PTO 05-2082

CY=JA DATE=19990511 KIND=A
PN=11-123080

GENE FOR RAFFINOSE SYNTHETASE, PRODUCTION OF RAFFINOSE, AND TRANSFORMED PLANT

[Rafinoosu gosei koso idenshi, rafinoosu no seizo-ho oyobi keishitsu tenkan shokubutsu]

Chieko Osumi, et al.

UNITED STATES PATENT AND TRADEMARK OFFICE
Washington, D.C. March 2005

Translated by: FLS, Inc.

PUBLICATION COUNTRY	(10):	JP
DOCUMENT NUMBER	(11):	11123080
DOCUMENT KIND	(12):	A
PUBLICATION DATE	(43):	19990511
PUBLICATION DATE	(45):	
APPLICATION NUMBER	(21):	09292969
APPLICATION DATE	(22):	19971024
ADDITION TO	(61):	
INTERNATIONAL CLASSIFICATION	(51):	C12N 15/09; A01H 5/00; C12N 9/00; //(C12N 9/00; C12R 1:19)
DOMESTIC CLASSIFICATION	(52):	
PRIORITY COUNTRY	(33):	
PRIORITY NUMBER	(31):	
PRIORITY DATE	(32):	
INVENTOR	(72):	OSUMI, CHIEKO; NOZAKI, YASUSHI; KIDA, TAKAO
APPLICANT	(71):	AJINOMOTO CO., INC.
TITLE	(54):	GENE FOR RAFFINOSE SYNTHETASE, PRODUCTION OF RAFFINOSE, AND TRANSFORMED PLANT
FOREIGN TITLE	[54A]:	RAFINOOSU GOSEI KOSO IDENSHI, RAFINOOSU NO SEIZO-HO OYOBI KEISHITSU TENKAN SHOKUBUTSU

[Title of the Invention] /1

GENE FOR RAFFINOSE SYNTHETASE, PRODUCTION OF RAFFINOSE, AND TRANSFORMED PLANT

[Claims] /

[Claim 1] Raffinose synthetase having activity for producing raffinose from sucrose and galactinol.

[Claim 2] Raffinose synthetase which is a protein indicated by the following (A), (B), (C), or (D):

(A) protein having the amino acid sequence expressed by sequence number 5 in the sequence charts

(B) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 5 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

[Claim 3] Raffinose synthetase according to Claim 1, having the following properties:

(1) action and substrate specificity: produces raffinose from sucrose and galactinol

(2) optimum pH: approximately 6 to 8

(3) optimum temperature: approximately 35°C to 40°C

(4) molecular weight:

(i) molecular weight measured by gel filtration chromatography: approximately 75 kDa to 95 kDa

---

*Numbers in the margin indicate pagination in the foreign text.

(ii) molecular weight measured by polyacrylamide gel electrophoresis: approximately 90 kDa to 100 kDa

(iii) molecular weight measured by SDS-polyacrylamide gel electrophoresis under reducing conditions: approximately 90 kDa to 100 kDa

(5) inhibition: inhibited by iodoacetamide, N-ethylmaleimide, and myoinositol

[Claim 4] Raffinose synthetase according to Claim 1, wherein the amino acid sequence contains the amino acid sequences expressed by sequence numbers 28 to 30 in the sequence charts.

[Claim 5] Raffinose synthetase which is a protein indicated by the following (C) or (D):

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

[Claim 6] Production of raffinose, characterized by producing raffinose by reacting a raffinose synthetase according to any one of Claims 1 to 5 with sucrose and galactinol.

[Claim 7] DNA coding a raffinose synthetase according to any one of Claims 1 to 5.

[Claim 8] DNA coding a protein indicated by the following (A), (B), (C), or (D):

(A) protein having the amino acid sequence expressed by sequence number 5 in the sequence charts

(B) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 5 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24

in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

[Claim 9] DNA coding a protein indicated by the following (C) or (D):

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

[Claim 10] DNA according to Claim 8, which is a DNA indicated by the following (a), (b), (c), or (d):

(a) DNA containing a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts

(b) DNA coding a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts, and a protein having activity for producing raffinose from sucrose and galactinol

(c) DNA containing a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts

(d) DNA coding a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts, and a protein having activity for producing raffinose from sucrose and galactinol

[Claim 10] DNA according to Claim 9, which is a DNA indicated by the following (c) or (d):

(c) DNA containing a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts /

(d) DNA coding a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts, and a protein having activity for producing raffinose from sucrose and galactinol

[Claim 12] DNA indicated by the following (e) or (f):

(e) DNA for hybridizing a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts or a complementary base sequence thereof under strict conditions

(f) DNA for hybridizing a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts or a complementary base sequence thereof under strict conditions

[Claim 13] Chimera gene containing the gene for raffinose synthetase or a part thereof, and a transcription control region which can be manifested in plant cells.

[Claim 14] Chimera gene according to Claim 1, wherein the above-mentioned gene for raffinose synthetase is a DNA according to any one of Claims 7 to 11.

[Claim 15] Chimera gene according to Claim 13 or 14, wherein the above-mentioned transcription control region is linked to the above-mentioned DNA so as to manifest an antisense RNA having a complementary sequence to the code strand of the above-mentioned DNA.

[Claim 16] Plant transformed by a chimera gene according to any one of Claims 13 to 15.

[Claim 17] Method for changing the raffinose family oligosaccharide content of the above-mentioned plant by transforming a plant by a chimera gene according to any one of Claims 13 to 15, and manifesting this gene in the plant cells.

[Detailed Explanation of the Invention]

[Industrial Field of Application]

The present invention pertains to raffinose synthetase, methods for synthesizing raffinose using raffinose synthetase or cell-derived products containing raffinose synthetase, DNA coding raffinose synthetase, and uses of this DNA in plants. Raffinose has an activity which propagates Lactobacillus bifidus, and is used in a variety of fields as a food ingredient or as pharmaceuticals such as organ preservative solutions.

[Prior Art]

Raffinose is a raffinose family oligosaccharide having a sucrose glycosyl group bonded to the α-1,6 position. Besides raffinose, raffinose family oligosaccharides include stachyose, which has two bonded galactose,

and verbascose, which has three bonded galactose. These saccharides are widely present as storage sugars in plant seeds such as bean seeds, rapeseeds, or cottonseeds and translocation sugars found in cucurbit plants such as cucumbers or melons, and in plants which have acquired cold resistance, such as rosette leaves or sugar beets.

Raffinose family oligosaccharides are biosynthesized as follows:

UDP-galactose + myoinositol → galactinol + UDP ... (a)
galactinol + sucrose → raffinose + myoinositol ... (b)
galactinol + raffinose → stachyose + myoinositol ... (c)

These reactions are catalyzed by (a) galactinol synthetase (GS: EC 2.4.1.123), (b) raffinose synthetase (RS: EC 2.4.1.82), and (c) stachyose synthetase (STS: EC 2.4.1.67)

Raffinose is extracted today from sugar beets, then isolated and refined by a sucrose refining process. However, raffinose reduces the crystallinity of sucrose. Therefore, sugar beets have been bred and improved to have low raffinose, producing sugar beets with a low raffinose content of 0.03% to 0.16% (*Enzyme Microb. Technol.*, Vol. 4 May, 130-135 (1982)). As a result, it is not easy to obtain raffinose effectively from sugar beets with such a low raffinose content.

As noted above, raffinose is found in the mature seeds of the bean family, starting with soybeans, and also in cucurbit plants such as cucumbers and melons. Mature soybean seeds contain the soybean oligosaccharides of sucrose (about 5% content), stachyose (about 4% content), and raffinose (about 1% content). These soybean oligosaccharides are recovered in a deproteinized fraction from defatted soybeans and condensed, then used in functional foods or the like. The yield of raffinose, however, is only 10% of all oligosaccharides.

An enzymatic synthesis method for raffinose has also been reported (*Trends in Glycoscience and Glycotechnology*, 7.34, 149-158 (1995)). This method synthesizes raffinose by condensing α-galactosidase to synthesize galactobiose, then converting to sucrose by galactosyl conversion using this galactobiose as a galactosyl group donor. This reaction, however, synthesizes 350 g galactobiose from 1.9 kg of lactose hydrolysate and gives 100 g raffinose from 190 g galactobiose and 760 g sucrose. Thus, this reaction has a low yield of raffinose, making it an ineffective synthesis method.

Besides methods such as discussed above, methods for breeding plants with a high content of raffinose by transforming enzyme genes in a biosynthesis system may be considered. For example, Kerr et al. cloned a gene for galactinol synthetase and used this gene to transform rapeseeds (WO93/02196). Although resulting in increased GS activity, however, this conversely reduced raffinose family oligosaccharides. Thus, this cloning

did not achieve the intended increase in biosynthesis of raffinose family oligosaccharides by introducing galactinol synthetase. Hence, no method has yet been offered for increasing the raffinose family oligosaccharide content in plants.

Raffinose family oligosaccharides have also been deliberately reduced. As noted earlier, raffinose family oligosaccharides are widely present mainly as storage sugars in plant seeds such as bean seeds, rapeseeds, or cottonseeds and translocation sugars found in cucurbit plants such as cucumbers or melons, and in plants which have acquired cold resistance, such as rosette leaves or sugar beets. Meals expressed from seeds such as cottonseeds also contain these raffinose family oligosaccharides. Most of these meals are used as animal feed, but humans or animals who lack α-galactosidase cannot directly consume raffinose family oligosaccharides. Furthermore, raffinose family oligosaccharides are known to reduce the metabolic energy efficiency of feeds by, for example, providing a source producing gas by intestinal bacteria, and removing raffinose family oligosaccharides from feed reportedly increases feeding efficiency in birds (Coon, *Proceeding Soybean Utilization Alternatives*, University of Minnesota, 203-211 (1989)). This has led to a demand for feed crops such as soybeans, rapeseeds, or cottonseeds having reduced raffinose family oligosaccharides.

Some of these plants have also been bred to increase oil content. Photosynthesis products are distributed in fats and oils, proteins, and saccharides, including raffinose family oligosaccharides. It has been reported that the oil and fat content is inversely correlated with the saccharide content in soybeans. Inhibiting production of raffinose family oligosaccharides is expected to increase the oil and fat content of soybeans having the same biosynthesis capacity.

From the standpoints discussed above, Kerr et al. reported developing a low raffinose family oligosaccharide soybean variety with 80% to 90% reduced raffinose family oligosaccharides by cross-breeding (WO93/00742). This is a developed variety, however, and cannot be applied to other varieties developed for purposes such as suitability for cultivation or disease resistance. This method also cannot be applied widely to different types of plants.

The raffinose contained in plants such as sugar beets or sugar cane is known to reduce sugar crystallinity. Therefore, these plants cannot be expected to have increased sugar production efficiency unless raffinose production is decreased, but no sugar beet has been developed which contains no raffinose.

As discussed above, previous refined raffinose synthetase has only been checked for enzymatic activity, and the enzyme has not been identified. This activity, moreover, has been low. Hence, there is a demand for raffinose synthetase with higher activity. Previous production of

raffinose also has low efficiency. Hence, there is a demand for production of raffinose with higher efficiency. On the other hand, there is a demand for breeding plants with reduced raffinose family oligosaccharides.
[Problems that the Invention is to Solve]

The problems addressed by the present invention are to acquire raffinose synthetase with high activity and the DNA coding this, and to offer an efficient method for enzymatic production of raffinose and methods for using plants with DNA coding raffinose synthetase.

[Means of Solving the Problems]

As a result of extensive studies toward solving the problems given above, the present inventors succeeded in refining raffinose synthetase from cucumber. The present inventors also conducted extensive studies to clone the gene coding this raffinose synthetase. They chemically synthesized a single-stranded DNA based on a base sequence hypothesized from the amino acid sequence of a peptide fragment of cucumber raffinose synthetase. Using this single-stranded DNA as a primer, they performed PCR using cDNA constructed from poly(A)$^+$ RNA extracted from cucumber as the matrix. As a result, they obtained a specific DNA fragment of the gene for raffinose synthetase. Furthermore, they hybridized a library of cucumber-derived cDNA using this DNA fragment as a probe, searched for a method for isolating the gene for raffinose synthetase, and isolated the gene for raffinose synthetase. They also conducted extensive studies to clone a gene for soybean-derived raffinose synthetase based on the information of the gene for cucumber-derived raffinose synthetase. As a result, they isolated a gene for soybean-derived raffinose synthetase. Using this isolated raffinose synthetase gene fragment, they constructed a chimera gene having a control region which can be manifested in plants. Furthermore, they were able to breed plants in which raffinose family oligosaccharides had been reduced by an introduced gene for raffinose synthetase.

Specifically, the present invention offers raffinose synthetase having activity for producing raffinose from sucrose and galactinol. In a preferred mode, the present invention offers raffinose synthetase which is a protein indicated by the following (A), (B), (C), or (D):

(A) protein having the amino acid sequence expressed by sequence number 5 in the sequence charts

(B) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 5 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

(C) protein having the amino acid sequence expressed by sequence number

24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

The present invention also offers raffinose synthetase having the following properties:

(1) action and substrate specificity: produces raffinose from sucrose and galactinol

(2) optimum pH: approximately 6 to 8

(3) optimum temperature: approximately 35°C to 40°C

(4) molecular weight:

(i) molecular weight measured by gel filtration chromatography: approximately 75 kDa to 95 kDa

(ii) molecular weight measured by polyacrylamide gel electrophoresis (Native PAGE): approximately 90 kDa to 100 kDa

(iii) molecular weight measured by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions: approximately 90 kDa to 100 kDa

(5) inhibition: inhibited by iodoacetamide, N-ethylmaleimide, and myoinositol

The present invention offers raffinose synthetase in which the amino acid sequence contains the amino acid sequences expressed by sequence numbers 28 to 30 in the sequence charts as a specific mode of the above-mentioned raffinose synthetase.

The present invention also offers raffinose synthetase which is a protein indicated by the following (C) or (D):

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

The present invention also offers production of raffinose,

characterized by producing raffinose by reacting the above-mentioned raffinose synthetase with sucrose and galactinol.

Furthermore, the present invention also offers DNA coding the above-mentioned raffinose synthetase, especially DNA coding a protein indicated by the following (A), (B), (C), or (D):

(A) protein having the amino acid sequence expressed by sequence number 5 in the sequence charts

(B) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 5 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

The present invention offers DNA indicated by the following (a), (b), (c), or (d) as specific modes of the above-mentioned DNA:
(a) DNA containing a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts

(b) DNA coding a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts, and a protein having activity for producing raffinose from sucrose and galactinol

(c) DNA containing a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts

(d) DNA coding a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts, and a protein having activity for producing raffinose from sucrose and galactinol

The present invention offers DNA which can be used for manifesting the antisense RNA or sense RNA of raffinose synthetase; that is, DNA indicated by the following (e) or (f):

(e) DNA for hybridizing a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts or a complementary base sequence thereof under strict conditions
(f) DNA for hybridizing a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts or a complementary base sequence thereof under strict conditions

Furthermore, the present invention offers a chimera gene containing the gene for raffinose synthetase or a part thereof and a transcription control region which can be manifested in plant cells, and a plant transformed by this chimera gene.

The present invention also offers a method for changing the raffinose family oligosaccharide content of the above-mentioned plant by transforming a plant by the above-mentioned chimera gene and manifesting this gene in the plant cells.

Hereafter, raffinose synthetase with the above-mentioned properties (1) to (5) or raffinose synthetase which is the above-mentioned protein (A), (B), (C), or (D) will be called simply "raffinose synthetase." Likewise, a DNA coding raffinose synthetase or a DNA coding raffinose synthetase and containing a non-translation region will be called a "gene for raffinose synthetase."

[Ideal Modes for Reducing the Invention to Practice]

Next, the present invention will be discussed in detail.

<1> Raffinose Synthetase of the Present Invention

The raffinose synthetase of the present invention is a protein indicated by the following (A), (B), (C), or (D):

(A) protein having the amino acid sequence expressed by sequence number 5 in the sequence charts

(B) protein comprised of an amino acid sequence in which one or more amino acids in the amino acid sequence expressed by sequence number 5 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactinol

(C) protein having the amino acid sequence expressed by sequence number 24 in the sequence charts

(D) protein comprised of an amino acid sequence in which one or more

amino acids in the amino acid sequence expressed by sequence number 24 in the sequence charts has been substituted, deleted, inserted, added, or inverted, and having activity for producing raffinose from sucrose and galactino

The raffinose synthetase of the present invention includes raffinose synthetase having the following properties:

(1) action and substrate specificity: produces raffinose from sucrose and galactinol

(2) optimum pH: approximately 6 to 8

(3) optimum temperature: approximately 35°C to 40°C

(4) molecular weight:

(i) molecular weight measured by gel filtration chromatography: approximately 75 kDa to 95 kDa

(ii) molecular weight measured by polyacrylamide gel electrophoresis: approximately 90 kDa to 100 kDa

(iii) molecular weight measured by SDS-polyacrylamide gel electrophoresis under reducing conditions: approximately 90 kDa to 100 kDa

(5) inhibition: inhibited by iodoacetamide, N-ethylmaleimide, and myoinositol

Raffinose synthetase with the properties recited above was isolated and refined from the main leaves of cucumber, and was first identified by the present inventors. As will be shown in the following working examples, this cucumber-derived raffinose synthetase contains the amino acid sequences expressed by sequence numbers 1 to 3 or 28 to 30 in the sequence charts within the amino acid sequences of its enzyme protein. The entire amino acid sequence of this raffinose synthetase is expressed by sequence number 5 in the sequence charts.

Raffinose synthetase is obtained from plants such as cucurbit plants, such as melons (Cucumis melo) or cucumbers (Cucumis sativus). The leaves, especially the leaf vein parts, and structures such as the seeds of these plants have an especially high content of raffinose synthetase.

Next, a method for isolating and refining raffinose synthetase from cucumber will be discussed as an example of production of the raffinose synthetase of the present invention. The leaf vein parts are collected from the main leaves of cucumbers six to ten weeks after planting, then crushed in a liquid nitrogen atmosphere using a mortar, for example, and

combined with a buffer to extract the protein. A substance may be added during this process to prevent degradation, loss, or other damage to raffinose synthetase; for example, a protease inhibitor such as PMSF (phenylmethanesulfonyl fluoride), or Polycral AT [as transliterated] (manufactured by Serva). Insoluble matter is removed from this extract by filtering or centrifuging, to give a raw extract.

Raffinose synthetase can be refined by combining and fractionating the raw extract obtained as described above by a conventional protein refining method; for example, anionic exchange chromatography, hydroxyapatite chromatography, gel filtration, or salting out.

Anionic exchange chromatography can be performed by using a column packed with a strongly acidic anionic exchanger such as HiTrapQ (manufactured by Pharmacia) or a weakly acidic anionic exchanger such as DEAE-TOYOPEARL (manufactured by Tosoh). The extract containing raffinose synthetase is passed through these columns to adsorb the enzyme to the columns. After washing the columns, the enzymes are extracted using a buffer with a high salt concentration. The salt concentration may be increased in stages to apply a concentration gradient during this process. For example, when using a HiTrapQ column, the raffinose synthetase adsorbed to the column is extracted by about 0.3 M NaCl. For DEAE-TOYOPEARL, the eluent is preferably a 0.05 M to 0.35 M NaCl concentration gradient, and for hydroxyapatite chromatography, the eluent is preferably a 0.01 M to 0.3 M phosphoric acid concentration gradient.

The order of the operations discussed above is not specially limited, and each operation may be repeated two or more times. Before passing a sample solution through each of these columns, the sample solution is preferably exchanged with a suitable buffer by a means such as dialysis. The sample solution may also be condensed at each of these stages.

Preferably, the activity of the raffinose synthetase contained in a fractionated fraction is assayed at each of the refining stages, and raffinose with the highest activity is collected and supplied to the next stage. An example of a method for assaying raffinose synthetase activity is the method using radioactive isotopes reported by H. Lehle et al. (*Eur. J. Biochem.*, 38, 103-110 (1973)). A variant of this method is to vary the reaction temperature and the substrate concentration. For example, 10 μL enzyme solution are combined with a reaction solution containing final concentrations of 10 mM^{14} C-sucrose, 20 mM galactinol, 25 mM HEPES (2-(4-(2-hydroxyethyl)-1-piperadinyl)ethanesulfonic acid)- NaOH (pH 7.0), and 5 mM DTT (dithiothreitol) to bring to 50 μL. This is reacted by incubating at 32°C for one hour, then combined with 200 μL ethanol and heated to 95°C for thirty seconds to stop the reaction. The centrifuged supernatant of this reaction solution is spotted on 3 MM Whatman filter paper and developed by a 4:1:2 ratio of n-propanol, ethyl acetate, and water. The uptake of ^{14}C into raffinose is investigated, and taken to show raffinose synthetase activity (nmol/time).

As an alternate method to the method described above, the present inventors developed the method of assaying raffinose synthetase activity by using HPLC (high-speed liquid chromatography) to determine the raffinose produced by raffinose synthesis. This method can assay activity more simply and faster than the method of H. Lehle et al., and is especially ideal for detecting active fractions in a refining operation.

Raffinose is synthesized by adding 10 to 50 μL raffinose synthetase solution to a reaction solution prepared to have a composition with the following final composition to bring to 100 μL, then reacting at 32°C for sixty minutes.

[Composition of Reaction Solution (Final Concentrations)]

2.5 mM	sucrose
5 mM	galactinol
5 mM	DTT
20 mM	tris hydrochloric acid buffer (pH 7.0)

After reacting as described above, the reaction solution is combined with four times its volume of ethanol and heated to 95°C for thirty seconds to stop the reaction. This is centrifuged, and the centrifuged supernatant is vacuum-dried, then dissolved in sterilized water. The raffinose in the reaction product is determined by HPLC to assay raffinose synthetase activity. The sugar analysis system DX500 (CarboPac PA1 column and pulsed amperometer detector (manufactured by Dionex)), for example, can be used to perform HPLC.

Fig. 1 shows results of assaying the raffinose produced by the method described above when reaction time was varied. As is clear from the graph, this method can assay raffinose synthetase activity simply and with good linearity.

Refined raffinose synthetase can be checked for purity and measured for molecular weight by a method such as gel electrophoresis or gel filtration chromatography. Enzymological properties may also be studied by assaying enzymatic activity while varying the temperature or pH of the reaction, or by adding various enzyme inhibitors, metallic ions, or the like to the reaction solution and assaying residual enzymatic activity. Furthermore, stable pH and temperature ranges may be investigated by assaying enzymatic activity after exposing raffinose synthetase to various pH conditions or temperature conditions for a set time.

Although the properties of the above-mentioned raffinose synthetase are determined in this way, it should be noted that different measurement conditions may give different results. For example, measurement of molecular weight by gel chromatography is influenced by the molecular weight marker or the type of buffer or gel filter medium used. Enzymatic activity also often differs, even at the same pH, depending on the type

of buffer or the salt concentration. Therefore, properties should be studied comprehensively rather than individually when identifying raffinose synthetase.

The raffinose synthetase of the present invention is obtained by isolating and refining from cucumber as described above, but the cucumber-derived, soybean-derived, or other plant-derived raffinose synthetases discussed below can be produced by introducing the DNA coding these into a host and manifesting the gene by methods conventionally used to produce other types of protein enzymes.

Although various procaryote cells, starting with Escherichia coli, and eucaryote cells, starting with Saccharomyces cerevisiae, may be considered for the host for manifesting the gene for raffinose synthetase, plant cells, especially cells derived from plants such as tobacco, cucumber, or Arabidopsis thaliana are preferred.

Recombinant plasmids used for transformation can be prepared by introducing a DNA coding raffinose synthetase into a manifest vector according to the type of cell that will manifest the gene. Plant manifest vectors have a promoter DNA sequence which is active in the plant or a combination of several of such promoter DNA sequences, and a terminator DNA which is active in the plant, and may have a sequence between these into which an extraneous gene can be introduced.

Examples of such promoters include promoters of genes manifesting the entire plant, such as the CaMV 35SRNA promoter, CaMV 19SRNA promoter, or nopalin synthetase promoter, promoters of genes manifesting green tissues, such as the Rubis CO small unit promoter, and promoters of genes manifesting specific parts such as seeds, such as napin or phaseolin. Examples of terminators such as described above include the nopalin synthetase terminator and Rubis CO small subunit 3'-side site.

Manifest vectors for plants are available commercially, and may be used; for example, pBI121, p35S-GFP (manufactured by CLONTECH). Vectors manifesting virus RNA may be used, and the genes such as epithelial proteins coded by these may be substituted for genes for raffinose synthetase.

Any method conventionally used for transformation may be used depending on the host cell to be supplied; for example, the Agrobacterium method, the particle-gun method, electroporation, or PEG.

"Genes coding cucumber-derived raffinose synthetase" include all such genes so long as they have raffinose synthetase activity when manifested, but preferred examples are genes which have a DNA coding the amino acid sequence expressed by sequence number 5 in the sequence charts, or genes having a base sequence expressed by sequence number 4 in the sequence charts. "Genes coding soybean-derived raffinose synthetase" include all such genes so long as they have raffinose synthetase activity

when manifested, but preferred examples are genes which have a DNA coding the amino acid sequence expressed by sequence number 24 in the sequence charts, or genes having a base sequence expressed by sequence number 23 in the sequence charts. Moreover, "genes coding the amino acid sequence expressed by sequence number 5 or 24 in the sequence charts" include a number of base sequences which are considered codon degenerates. That is, a base sequence may be selected from several such base sequences by considering the elements of a gene manifestation system, such as avoiding codons which take priority depending, for example, on the type of host cell, or higher structures formed by transcribed RNA. The selected base sequence may be a DNA which has been cloned from nature, or a DNA which has been chemically synthesized artificially.

<2> DNA Coding Raffinose Synthetase of the Present Invention

A DNA coding raffinose synthetase can be acquired by preparing a cDNA library from poly(A)$^+$ RNA isolated from plants such as cucumber, then screening this cDNA library by hybridization. The probe used during hybridization can be acquired by amplification using an oligonucleotide synthesized based on a partial amino acid sequence of the raffinose synthetase protein as a promoter.

Next, a specific example of a method for acquiring DNA of the present invention from poly(A)$^+$ RNA derived from cucumber will be discussed. The poly(A)$^+$ RNA extraction site may be any part of the cucumber plant so long as the site manifests the gene for raffinose synthetase. Poly(A)$^+$ RNA can be obtained, for example, from the leaves, stems, buds, fruit, or seeds at various growth stages. Preferably, however, the material used is the leaves developing after fruiting, especially the leaf vein parts.

The method for extracting the total RNA from cucumber tissue is not limited so long as the method can obtain RNA efficiently with little damage. For example, a conventional method such as the phenol/SDS method or the guanidine isothiocyanate/cesium chloride method can be used. Poly(A)$^+$ RNA can be isolated from the total RNA obtained in this way using an oligo(dT) carrier. A kit which can obtain poly(A)$^+$ RNA without extracting the total RNA (such as MPG Direct mRNA Purification Kig, CPG, Inc.) may also be used.

A DNA fragment of the probe to be used for screening a cDNA library can be obtained by PCR. A single-stranded DNA having a base sequence hypothesized from the amino acid sequence of an already known peptide fragment, such as the amino acid sequence expressed by sequence numbers 1 to 3 in the sequence charts, is chemically synthesized, then PCR is performed using this as a primer. Any part of the amino acid sequence of the already known peptide fragment may be used for the primer, but preferably, a base sequence is selected which has few codon degenerates and does not form complicated higher structures. RACE (Rapid Amplification of cDNA End: *PCR PROTOCOLS, A Guide to Methods and Applications*, Academic

Press, Inc., pp. 28-38) may also be performed.

The matrix of this PCR is preferably a cDNA library and a single-stranded cDNA. Poly(A)$^+$ RNA, and in some cases the total RNA, may be used when using a heat-resistant DNA polymerase which has reverse transcription enzymatic activity during PCR reactions.

To construct a cDNA library, first, a single-stranded DNA is synthesized by a reverse transcription enzyme using a primer such as an oligo(dT) primer or a random primer with poly(A)$^+$ RNA as a matrix, then a double-stranded DNA is synthesized by a method such as the Gubler-Hoffman method or the Okayama-Berg method (*Molecular Cloning*, 2nd edition, Cold Spring Harbor Press, 1989). If too little gene for raffinose synthetase is manifested, the cDNA may be amplified by PCR using a cDNA library construction kit (such as Capfinder PCR cDNA Library Construction Kit (CLONTECH)). A cDNA synthesized in this way can be cloned to a cloning vector such as a phage vector or a plasmid by a means such as smoothing the terminals, adding a linker, or adding a restriction enzyme site by PCR.

From DNA fragments obtained by PCR as described above, a part characterizing the cDNA of raffinose synthetase is selected as a probe for hybridization. Preferably, a DNA fragment near the 5′ terminal is selected. The amplified DNA fragment selected in this way is refined from the PCR reaction solution. During this refining, the amplified DNA fragment may be sub-cloned using a plasmid, and after a large amount of the plasmid has been prepared, the plasmid may be cut by a restriction enzyme and refined by electrophoresing, then cutting out the gel. Alternately, the intended site alone may be amplified and used by performing PCR using a plasmid as a matrix. Furthermore, if the initial amount of amplified DNA fragment is enough, the amplified DNA fragment may be electrophoresed without sub-cloning, and a band of the intended DNA fragment may be refined from the gel fragment containing this band.

Screening to obtain an intended clone from a cDNA library is done by hybridization. A DNA fragment obtained by the method described above can be made a probe for hybridization by labeling. A number of means can be used for labeling, such as a radioisotope or biotin, but labeling by the random priming method is preferred. Screening may also be done using PCR instead of hybridization. Furthermore, hybridization and PCR may be combined. Sequence number 4 in the sequence charts shows an example of the base sequence of a DNA coding cucumber-derived raffinose synthetase obtained as described above and the amino acid sequence hypothesized from this base sequence. Sequence number 5 shows this amino acid sequence alone. Transformant AJ13263 of Escherichia coli JM109 supporting plasmid pMossloxCRS, which contains a DNA fragment containing a DNA coding raffinose synthetase obtained in Working Example 3 discussed below, was deposited for international safekeeping based on the Budapest Convention with the National Institute of Bioscience and Human Technology, Agency

of Industrial Science and Technology, Ministry of International Trade and Industry (Post Office Box 305, 1-1-3 Higashi, Tsukuba-shi, Ibaraki Prefecture, Japan) on November 19, 1996, and assigned the receipt number of FERM BP-574.

Furthermore, genes for raffinose synthetase from other plants can be acquired using the gene for raffinose synthetase obtained from one plant as described above. As noted above, the plant from which raffinose synthetase is to be obtained may be any plant which synthesizes raffinose synthetase, such as soybean, fava bean, rapeseed, sunflower, cotton, or sugar beet. As an example, acquisition of a DNA coding soybean raffinose synthetase using a DNA coding cucumber-derived raffinose synthetase will be discussed.

A gene for soybean raffinose synthetase can be acquired by preparing a cDNA library from poly(A)$^+$ RNA derived from soybean, then cloning a cDNA library using a probe selected based on the base sequence of a DNA coding cucumber-derived raffinose synthetase.

The RNA extraction site may be any part of the soybean plant so long as the part manifests the gene for raffinose synthetase, but is preferably seeds, especially immature seeds after blossoming, when the soybean plant is producing raffinose family oligosaccharides.

The method for extracting the total RNA from immature soybean seeds is not specially limited so long as the method obtains RNA efficiently and with little damage. Any of the methods discussed above for cucumber can be used.

The hybridization probe must have a sequence that is highly homologous with the gene for soybean-derived raffinose synthetase. Although the probe used during hybridization can be the gene for cucumber-derived raffinose synthetase, it is better to probe this gene for the region which preserves raffinose synthetase. It is not possible to specify such a sequence, however, from the information of the gene for cucumber-derived raffinose synthetase. Obtaining a hybridization probe with such a sequence requires one of several methods such as the following: A fragment obtained by cutting a gene for cucumber raffinose synthetase by a suitable restriction enzyme may simply be Northern-hybridized with soybean RNA to probe for a DNA fragment for hybridization. A DNA fragment can also be obtained by RT-PCR with soybean RNA as a matrix using a primer synthesized based on the amino acid sequence of cucumber-derived raffinose synthetase. A DNA fragment can also be obtained by subjecting Arabidopsis RNA to RT-PCR using an oligonucleotide synthesized based on an Arabidopsis sequence which is homologous with DNA coding cucumber-derived raffinose synthetase as a primer.

Preferably, a fragment which is highly homologous with the intended gene is obtained as follows: First, a software such as Genetix Mac is

used to search the GenBank database for an EST sequence such as Arabidopsis which is homologous with a gene for cucumber-derived raffinose synthetase. The region of the resulting sequence which is highly homologous with the gene for cucumber-derived raffinose synthetase is thought to contain a sequence which is preserved between raffinose synthetases derived from different plants. To obtain a DNA fragment of this region, for example, an oligonucleotide which has been synthesized based on the sequence of the highly homologous region can be amplified by PCR using a single-stranded DNA prepared from the RNA of Arabidopsis as a matrix. The base sequence of this amplified fragment is cut, and a sequence which is highly homologous with cucumber is selected. The resulting DNA fragment is labeled as described above, then used as a probe.

A cDNA library is screened to obtain an intended clone by hybridizing in the same way as when cloning the cucumber gene. Sequence number 23 in the sequence charts shows the base sequence of a DNA coding soybean-derived raffinose synthetase obtained as described above and the amino acid sequence hypothesized from this base sequence. Sequence number 24 shows this amino acid sequence alone. Maximum matching considering /1
gaps indicated 38% homology in the amino acid sequence and 50% homology in the base sequence of the soybean-derived raffinose synthetase with the cucumber-derived raffinose synthetase acquired by the method described above. Transformant AJ13379 of Escherichia coli JM109 supporting plasmid pMossloxSRS, which contains a DNA fragment containing a DNA coding raffinose synthetase obtained in Working Example 4 discussed below, was deposited for international safekeeping based on the Budapest Convention with the National Institute of Bioscience and Human Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry (Post Office Box 305, 1-1-3 Higashi, Tsukuba-shi, Ibaraki Prefecture, Japan) on October 20, 1996, and assigned the receipt number of FERM BP-6149.

Genes for raffinose synthetase from other plants can be acquired using the information of cucumber-derived raffinose synthetase and soybean-derived raffinose synthetase. Single-stranded DNA having a base sequence hypothesized from the amino acid sequence preserved in both proteins, such as sequence number 28 (amino acid numbers 199 to 208 of sequence number 24), 29 (amino acid numbers 302 to 314 of sequence number 24), or 30 (amino acid numbers 513 to 527 of sequence number 24) in the sequence charts, are synthesized, and RT-PCR is performed using these as primers. Although any of the sequences discussed above may be used for primers, preferably, sequences thought to have few codon degenerates and to not form complicated higher structures are selected. A cDNA is synthesized from the total RNA, or in some cases the poly(A)$^{+}$ RNA, of the plant from which a gene is to be acquired, and PCR is performed using this as a matrix. The resulting DNA fragment is cloned to a suitable vector, and the base sequence is analyzed. It may be confirmed whether this base sequence is homologous with the gene for cucumber- or soybean-derived raffinose synthetase, or whether the translated amino acid sequence is

homologous with the gene for cucumber- or soybean-derived raffinose synthetase. The DNA fragment obtained as a result can be used for screening a cDNA library. RACE may also be performed using a single-stranded cDNA synthesized from the total RNA, or in some cases the poly$(A)^+$ RNA, of the plant from which a gene is to be acquired as a matrix.

The DNA of the present invention may code a raffinose synthetase protein which has one or several substituted, deleted, inserted, added, or inverted amino acids in one or more positions, so long as this does not damage the activity of the raffinose synthetase to be coded; that is, its activity producing raffinose from sucrose and galactinol. How many this "several" is differs depending on the type and position of the amino acid residual group in the stereo structure of the protein. This is because some amino acids, such as isoleucine and valine, have high affinity and the differences between such amino acids do not greatly affect the stereo structure of the protein. Therefore, it is enough to have at least 35% to 40% homology with all 784 amino acid residual groups comprising cucumber-derived raffinose synthetase, and have raffinose synthetase activity. Furthermore, preferably, this DNA has at least 65% homology from the 510^{th} amino acid to the 610^{th} amino acid, and more preferably, the "several" amino acids are two to forty, preferably two to twenty, and more preferably two to ten amino acids. Likewise, it is enough to have at least 35% to 40% homology with all 784 amino acid residual groups comprising soybean-derived raffinose synthetase, and have raffinose synthetase activity. Furthermore, preferably, this DNA has at least 65% homology from the 478^{th} amino acid to the 577^{th} amino acid. "Homology" here refers to the maximum matching considering gaps.

This includes genes having at least about 50% homology along the total length of the gene and at least 65% homology over about 300 bases within this total length. The base sequence information of such a gene can be obtained by searching a database such as GenBank for a gene which is highly homologous with a gene for cucumber-derived raffinose synthetase. A program such as GENETIX-MAC (a gene information processing software manufactured by the Software Development Co.) employing the Lipman-Pearson method may be used for the homology analysis program. Programs published on the Internet may also be used. Some of the base sequences obtained by such methods contain the total length of the gene, and some do not. If a base sequence does not contain the total length of the gene, the total length of the gene can be easily acquired by the 5' RACE method or the 3' RACE method with an RNA extracted from the intended plant tissue as a matrix by using a primer corresponding to the site which is highly homologous with a gene for cucumber-derived raffinose synthetase. The total length of the resulting gene may be recombined as discussed early with a suitable vector offered by a kit such as Soluble Protein Expression System (INVITROGEN), Tight Control Expression System (INVITROGEN), or QIAexpress System (QIAGEN) to manifest the gene. After assaying raffinose synthetase activity by the method described below, a clone showing activity may be selected. Gene manifestation methods are

discussed in detail by sources such as *Plant Molecular Biology, A Laboratory Manual* (Melody S. Clark (ed.), Springer).

A DNA coding substantially the same protein as this raffinose synthetase is obtained by transforming a base sequence having substituted, deleted, inserted, added, or inverted amino acids in specific positions using a method such as site-specific transformation. A DNA which has been transformed in this way may also be obtained by conventional /1 mutation processing. Examples of mutation processing are methods for *in vitro* processing of a DNA coding raffinose synthetase by hydroxylamine or the like, and methods for processing an Escherichia bacterium supporting a DNA coding raffinose synthetase by a mutating agent commonly used in artificial mutation, such as ultraviolet radiation or N-methyl-N'-nitro-N-nitrosoguanidine (NTG).

Substitution, deletion, insertion, addition, or inversion of bases such as discussed above also includes naturally occurring mutations such as occur due to differences in morphology or species of cucumbers or soybeans, multiple copies of genes, or differences in organs or tissues.

A DNA coding substantially the same protein as raffinose synthetase is obtained by manifesting a DNA having a mutation such as described above by a suitable cell, and investigating the raffinose synthetase activity of the manifestation. A DNA coding substantially the same protein as raffinose synthetase may also be obtained by hybridizing under stringent conditions and isolating a DNA coding a protein having raffinose synthetase activity from a DNA coding raffinose synthetase with a mutation or cells supporting this DNA; for example, a DNA having a base sequence comprised of base numbers 56 to 207 of the base sequence expressed by sequence number 4 in the sequence charts, or a base sequence comprised of base numbers 156 to 2405 of the base sequence expressed by sequence number 23. "Stringent conditions" here means conditions which form so-called specific hybrids and do not form nonspecific hybrids. It is difficult to enumerate these conditions in numerical terms, but to give a examples, these are conditions which hybridize highly homologous DNA, such as DNA having at least 50% homology, but do not hybridize DNA with less homology than this, or conventional Southern hybridization washing conditions, which are conditions which hybridize a salt concentration corresponding to 1 x SSC and 0.1% SDS, and preferably 0.1 x SSC and 0.1% SDS, at 60°C. Genes hybridized under such conditions include genes which have a stop codon midway and genes which have lost activity due to mutation of their active center, but these can easily be removed by linking to a commercial active manifest vector and assaying raffinose synthetase activity by the method described above.

Moreover, if a DNA of the present invention is to be used to manifest the antisense RNA of raffinose synthetase, this DNA need not code active raffinose synthetase. The sense RNA may also inhibit the intrinsic function of a gene. In this case, the DNA also need not code active raffinose

synthetase, and need not contain the total length of the gene so long as the translation region is about 500 base pairs having at least 60% homology. Examples of such DNA are the following DNA (e) or (f):

e) DNA for hybridizing a base sequence comprised of at least base numbers 56 to 2407 of the base sequence expressed by sequence number 4 in the sequence charts or a complementary base sequence thereof under strict conditions

(f) DNA for hybridizing a base sequence comprised of at least base numbers 156 to 2405 of the base sequence expressed by sequence number 23 in the sequence charts or a complementary base sequence thereof under strict conditions

The methods by which the present inventors successfully cloned the intended cDNA of cucumber-derived or soybean-derived raffinose synthetase have been discussed above. The following are examples of other methods:

(1) Cucumber- or soybean-derived raffinose synthetase is isolated and refined, and the entire base sequence is chemically synthesized based on the amino acid sequence to be determined or the amino acid sequence expressed by sequence number 5 or sequence number 24.

(2) A chromosome DNA is prepared from a cucumber or soybean plant, a chromosome DNA library is prepared using a plasmid vector or the like, and the gene for raffinose synthetase is acquired from this library by hybridization or PCR. Moreover, a chromosome-derived gene for raffinose synthetase may be expected to contain an intron in the coding region, but even a DNA cut by such an intron will contain DNA of the present invention so long as the DNA codes raffinose synthetase.

(3) Poly(A)$^+$ RNA is fractionated, for example, by molecular weight and supplied to an *in vitro* translation system using wheat germ or rabbit reticular leukocytes. A fraction containing mRNA coding a polypeptide having raffinose synthetase activity is determined, and the intended cDNA fragment is prepared and acquired from this fraction.

(4) An anti-cucumber raffinose synthetase antibody or an anti-soybean raffinose synthetase antibody is prepared. A cDNA library is set in a protein manifest vector, and the protein coding cDNA is manifested by infecting a suitable host. The intended cDNA is screened using the previous antibody.

(5) A suitable primer is synthesized from the amino acid sequence of a peptide fragment, the sequence containing the terminal is amplified by RACE, and this sequence is cloned.

To manifest the gene for raffinose synthetase, the DNA of the region coding the enzyme may be recombined with various manifest vectors to

manifest the gene. Details are given in sources such as *Plant Molecular Biology—A Laboratory Manual* (M. S. Clark (ed.), Springer). A commercial manifest vector may be used as the vector. Manifestation can be confirmed by assaying activity by the method described in this specification. /1

A method for manifesting raffinose synthetase activity by a gene for soybean-derived raffinose synthetase will be discussed as an example. An NdeI restriction enzyme site just above the 156^{th} base ATG and a BamHI site below the 2405^{th} base of the gene for raffinose synthetase contained in pMOSSloxSRS are added by PCR using a primer designed so as to contain both of these restriction enzyme sites. Next, the previous gene for raffinose synthetase refined in phenol chloroform and pET3a are consumed by NdeI and BamHI. Each of the consumed DNA is refined by agarose gel electrophoresis. The gene for soybean-derived raffinose synthetase has a BamHI site inside. Therefore, a fragment of the intended size may be selected by inserting a mutation beforehand by a means such as PCR, or by agarose electrophoresis. Furthermore, the base sequence is analyzed to confirm that the gene for raffinose synthetase starts from the ATG codon. This vector is transformed into E. coli BL21 (DE3) pLysE and selected by an LB medium containing chloramphenicol and ampicillin. PCR is used to confirm that the fragment has been introduced into the transformant, and the intended transformant is cultured in an LB medium to give a biomass. The biomass is incubated at 100°C for three minutes in a gel loading buffer containing SDS, and electrophoresed on an SDS polyacrylamide gel. A band of protein of the intended size is confirmed. The organism selected in this way is pulverized by a means such as sonics to extract the protein. The raffinose synthetase activity of this protein extract may be assayed by the method described in this specification.

<3> Production of Raffinose of the Present Invention

During production of raffinose of the present invention, raffinose is produced by reacting the raffinose synthetase described above with sucrose and galactinol. When raffinose synthetase reacts with sucrose and galactinol, the galactose residual group component of galactinol is transferred to the sucrose to produce raffinose. This transfer produces myoinositol, which is a component of galactinol.

The raffinose synthetase used when producing raffinose may be an enzyme extracted from a plant, or an enzyme produced by gene recombination using the DNA of the present invention.

Raffinose synthetase may be reacted with sucrose and galactinol by packing a column with a fixed enzyme or fixed cells prepared by fixing raffinose synthetase or a cell capable of producing raffinose synthetase to a carrier such as alginate gel or polyacrylamide gel, then pouring a solution containing sucrose and galactinol into this column. Materials and methods conventionally used in bioreactors can be employed to fix the raffinose synthetase or cells to a carrier.

For example, raffinose can be synthesized by adding raffinose synthetase to a solution containing sucrose and galactinol or a solution such as a buffer. The above solution is buffered to a pH in a range of about 6 to 8, and preferably close to pH 7. The reaction temperature is about 28°C to 42°C, preferably a range of 35°C to 40°C, and especially preferably about 38°C. Moreover, the raffinose synthetase of the present invention is stable in a range of about pH 5 to 8, and especially at close to pH 6. This enzyme is also stable in a temperature range no higher than about 40°C.

The enzymatic activity of raffinose synthetase of the present invention is blocked by iodoacetamide, N-ethylmaleimide, $MnCl_2$, $ZnCl_2$, and $NiCl_2$. Therefore, the reaction solution should not contain any of these substances.

The concentrations of galactinol and sucrose added to the reaction solution are ideally at least 5 mM galactinol and at least 1.5 mM sucrose. The weight of raffinose synthetase added to the reaction solution may be adjusted according to the weight of the substrate.

Gel filtration chromatography is an example of a method for separating raffinose from unreacted sucrose and galactinol contained in the reaction solution and myoinositol produced by the enzymatic reaction.

<4> Chimera Gene and Transformed Plant of the Present Invention

The chimera gene of the present invention contains the gene for raffinose synthetase or a part thereof, and a transcription control region which can be manifested in plant cells. An example of the gene for raffinose synthetase is the DNA coding raffinose synthetase of the present invention discussed in section <2>. Furthermore, if the chimera gene of the present invention is to be used as an antisense gene, a non-translation region of the gene for raffinose synthetase or a part thereof can sometimes be used besides a DNA coding raffinose synthetase. Examples are the sequence expressed by base numbers 1 to 55 (5′ non-translation region) or 2407 to 2517 (3′ non-translation region) in sequence number 4 in the sequence charts, and base numbers 1 to 155 or 2406 to 2765 in sequence number 23 in the sequence charts.

Linking the transcription control region in the chimera gene of the present invention to a DNA coding raffinose synthetase so as to manifest an mRNA (sense RNA) which is homologous to the code strand of this DNA causes the plant cell into which this chimera gene has been introduced to manifest raffinose synthetase and increases its content of raffinose family oligosaccharides. Linking the above-mentioned transcription control region to the above-mentioned DNA so as to manifest an RNA /1
with the complementary sequence to the above-mentioned code strand of DNA (antisense RNA) or so as to manifest a sense RNA for a fragment of

part of the gene for raffinose synthetase, preferably at least about 200 base pairs of the upper code region, on the other hand, causes the plant cell into which these chimera genes have been introduced to inhibit manifestation of intrinsic raffinose synthetase and reduces its content of raffinose family oligosaccharides.

Transforming a plant by a chimera gene of the present invention and manifesting this gene in plant cells as described above can change the raffinose family oligosaccharide content in the above-mentioned plant.

Examples of plants to which the present invention may be applied include soybeans, rapeseeds, and cotton, which are oil-producing plants, sugar beets and sugar cane, which are sugar-producing plants, and the model plant of Arabidopsis thaliana.

As noted earlier, examples of transcription control regions which can be manifested in plant cells include promoters of genes manifesting the entire plant, such as the CaMV 35SRNA promoter, CaMV 19SRNA promoter, or nopalin synthetase promoter, promoters of genes manifesting green tissues, such as the Rubis CO small unit promoter, and promoters of genes manifesting specific parts such as seeds, such as napin or phaseolin. Terminators such as the nopalin synthetase terminator or Rubis CO small subunit 3'-side site may be linked to the 3' terminal of a chimera gene.

Any conventional method for transforming a plant by a chimera gene may be used depending on the type of host cell to be supplied; for example, the Agrobacterium method, the particle-gun method, electroporation, or PEG.

Examples of transformation methods for introducing a chimera gene into a plant include the Agrobacterium method, the particle-gun method, electroporation, and PEG. A specific example of the Agrobacterium method is the method of using a binary vector. Specifically, a plant is infected by a vector containing T-DNA derived from a Ti plasmid, a replication starting point which is functional in microorganisms such as Escherichia coli, and a marker gene for selecting a plant cell or microorganism cell to support a vector. Seeds collected from this plant are sown, and a plant into which the vector has been introduced is selected as an index of manifestation of the marker gene. The intended transformed plant can be acquired by assaying the raffinose synthetase activity of the selected plant, or selecting a plant in which the raffinose family oligosaccharide content has changed.

Next, a method for introducing a chimera gene into soybeans will be discussed. Any of the particle-gun method (*Pro. Natl. Acad. Sci. USA*, 86, 145 (1989), *TIBTECH*, 8, 145 (1990), *Bio/Technology*, 6, 923 (1988), *Plant Physiol.*, 87, 671 (1988), *Develop. Genetics*, 11, 289 (1990), *Plant Cell Tissue & Organ Culture*, 33, 227 (1993)), the Agrobacterium method (*Plant Physiol.*, 91, 1212 (1989), WO94/02620, *Plant Mol. Biol.*, 9, 135 (1987), *Bio/Technology*, 6, 915 (1988)), or electrophoresis (*Plant

Physiol., 99, 81 (1992), *Plant Physiol.*, 84, 856 (1989), *Plant Cell Reports*, 10, 97 (1991)) can be used to transform soybeans.

The particle-gun method may use embryogenic tissue or the hypocotyl of immature seeds thirty to forty days after opening. About one gram of embryogenic tissue may be spread on a Petri dish, then particles such as gold or tungsten particles coated with the intended chimera gene may be driven in. One to two hours later, the tissue is transferred to a liquid medium and cultured. After two weeks, the tissue is transferred and cultured in a medium having an antibiotic introduced for selecting a transformant. After six weeks, a green resistant indeterminant embryo is obtained. This embryo is transferred to a fresh medium and cultured to regenerate the plant. Alternately, if using the hypocotyl, the hypocotyl is extracted under sterile conditions and treated by a particle gun, then cultured in an MS medium (Murashige and Skoog, *Physiologia Plantrum*, 15, 473-497 (1962)) containing a high concentration of cytokinin. After culturing for two weeks in the dark, the hypocotyl is cultured in an MS medium with a lower content of cytokinin for twelve to sixteen hours exposed to light. Preferably, an antibiotic for use as a selection marker has been added to the medium beforehand. Once a polyblast has been formed by the transplanted tissue, the polyblast is rooted by transferring to a medium with no hormone added. This rooted polyblast is then transferred to room temperature and cultivated.

Methods using Agrobacterium preferably use Cotyledonary nod as the plant tissue. Any commercial Agrobacterium such as LBA4404, C58, or Z707 can be used, but Z707 is preferred. A plasmid in which the intended gene has been introduced into pMON530 (Monsanto Co.), for example, can be used as a vector. The plasmid is introduced into Agrobacterium tumefaciens Z707 (Hepburn et al., *J. Gen. Microbiol.*, 131, 2961 (1985)) by a method such as the direct freeze-thaw method (An et al., *Plant Mol. Biol. Manual*, A3: 1-19, 1988). The Agrobacterium transformed by this chimera gene is cultured overnight, centrifuged at 5000 rpm for five minutes, then suspended in a B5 suspension medium. Soybean seeds are sterilized, then cultured in a 1/10 concentration B5 medium for three days to germinate. The cotyledon are cut off, then suspended in the Agrobacterium medium and cultured for two hours. This cotyledon is transferred to a B5 medium (Gamborg B5 medium (*Exp. Cell. Res.*, 50, 151 (1968) containing Gamborg Vitamin B5, 3% sucrose, 5 μM benzyl aminopurine, 10 μM IBA, and 100 /1
μM acetosyringon) and cultured for three days under conditions of 25°C and 23 hours of light exposure (60 $\mu Em^{-2}S^{-1}$). Next, this is cultured in a B5 medium (5 μM benzyl aminopurine, 100 mg/L carbenicillin, 100 mg/L vancomycin, 500 mg/L cefotaxime) for four days at 25°C while changing the medium daily to remove the Agrobacterium. Following this, this is cultured in a B5 medium (200 mg/L kanamycin). A multiple shoot forms in one to two months. This is cultured in a B5 culture (0.58 mg/L gibberellin, 50 mg/L kanamycin) to elongate the shoot. Next, this is transferred to a B5 medium (10 μM IBA) and rooted. A transformant can be obtained by habituating and cultivating the young rooted plant at room temperature.

A transformed plant having the gene for raffinose synthetase introduced can be easily confirmed by extracting DNA from the transformant and Southern-hybridizing this using the gene for raffinose synthetase as a probe.

[Working Examples]

Next, the present invention will be discussed in greater detail by working examples. First, the method for confirming active fractions during the refining steps and the method for assaying raffinose synthetase activity used to study enzyme characteristics in the following working examples will be discussed.

<Assaying Raffinose Synthetase Activity>

Raffinose synthetase activity was assayed by using HPLC (high speed liquid chromatography) to determine the raffinose produced by raffinose synthesis. The sugar analysis system DX500 (CarboPac PA1 column and pulsed amperometer detector (manufactured by Dionex)) was used to perform HPLC.

Raffinose was synthesized by adding 10 to 50 μL raffinose synthetase solution to a reaction solution prepared to have a composition with the following final compositions to bring to 100 μL, then reacting at 32°C for sixty minutes.

[Composition of Reaction Solution (Final Concentrations)]

2.5 mM	sucrose
5 mM	galactinol
5 mM	DTT
20 mM	tris hydrochloric acid buffer (pH 7.0)

After reacting as described above, the reaction solution was combined with four times its volume of ethanol and heated to 95°C for thirty seconds to stop the reaction. This was centrifuged, and the centrifuged supernatant was vacuum-dried, then dissolved in sterilized water. The raffinose in the reaction product was determined by HPLC to assay raffinose synthetase activity.

[Working Example 1] Refining Raffinose Synthetase from Cucumber

<1> Extraction of Raffinose Synthetase from Cucumber

Leaf vein parts were collected from the main leaves of cucumber (trade name "SUYOU") six to ten weeks after sowing, freeze-dried in liquid nitrogen, and stored at -80°C. About 200 g of freeze-dried leaf vein parts were crushed in liquid nitrogen using a mortar, then combined with Buffer 1 (40 mM tris hydrochloric acid buffer (pH 7.0), 5 mM DTT, 1 mM PMSF (phenylmethanesulfonyl fluoride), 1% Polycral AT; manufactured by Serva)

to extract the protein. The extract was filtered by a means such as gauze or Miracloth (Calbiochem-Novobiochem), and the filtrate was centrifuged for sixty minutes at 4°C and about 30,000 x g. The resulting centrifuge supernatant formed a raw extract.

<2> Anion Exchange Chromatography (1)

About 560 mL of raw extract obtained as described above was supplied to five connected columns of a strongly basic anion exchange chromatography column (HiTrapQ; manufactured by Pharmacia, 1.6 cm x 2.5 cm) balanced with Buffer 2 (20 mM tris hydrochloric acid buffer (pH 7.0), 5 mM DTT), and raffinose synthetase activity was adsorbed to the column. Next, the column was washed with five times the volume of the column of Buffer 3 (20 mM tris hydrochloric acid buffer (pH 7.0), 0.2 M NaCl, 5 mM DTT) to wash off unadsorbed proteins, then raffinose synthetase activity was eluted from the column by 50 mL of Buffer 4 (20 mM tris hydrochloric acid buffer (pH 7.0), 0.3 M NaCl, 5 mM DTT).

<3> Anion Exchange Chromatography (2)

About 75 mL of the extract described above were placed in a dialysis tube (Pormembranes MWCO: 10,000; manufactured by Spectra) and dialyzed overnight at 4°C against 10 L of Buffer 5 (20 mM tris hydrochloric acid buffer (pH 7.0), 0.05 M NaCl, 5 mM DTT). The dialyzed sample was supplied to a weakly basic anion exchange chromatography column (DEAE-TOYOPEARL; manufactured by Tosoh, 2.2 x 20 cm) balanced with Buffer 5, and raffinose synthetase activity was adsorbed to the column. Next, the column was washed with five times the volume of the column of Buffer 5 to wash off unadsorbed proteins, then enzymatic activity was eluted from the column by applying a linear gradient of 0.05 M to 0.35 M concentrations of NaCl against twenty times the column volume.

<4> Gel Filtration Chromatography

About 160 mL of extract obtained as described above were condensed to 6.5 mL using a condenser (Centriprep 10, manufactured by Amicon). This condensate was supplied 3 mL at a time to a gel filtration chromatogram (Superdex 200 pg; manufactured by Pharmacia, 2.6 cm x 60 cm). The column was balanced and eluted using Buffer 6 (20 mM tris hydrochloric acid buffer (pH 7.0), 0.1 M NaCl, 5 mM DTT, 0.02% Tween 20). Fractions with /1
raffinose synthetase activity were collected from the fractionated fractions.

<5> Hydroxyapatite Chromatography

About 25 mL of fractions with raffinose synthetase activity fractionated by gel filtration were condensed by a Centriprep 10. The buffer was replaced using Buffer 7 (0.01 M sodium phosphate buffer (pH 7.0), 5 mM DTT, 0.02% Tween 20). About 1.2 mL of the resulting condensate

were supplied to a hydroxyapatite column (Bio-Scale CHT-1; manufactured by BioRad, 0.7 x 5.2) balanced beforehand with the same buffer, and raffinose synthetase activity was adsorbed to the column. The column was washed with five times the volume of the column (10 mL) of the same buffer, then fractionated by applying a linear gradient of 0.01 M to 0.3 M concentrations of phosphoric acid against twenty times the column volume.

<6> Hydroxyapatite Rechromatography

The active fractions obtained by hydroxyapatite chromatography as described above were subjected to rechromatography to produce a refined raffinose synthetase fraction (about 2 mL).

The weight of protein in this active fraction was about 200 mg. The total activity was 5700 nmol/h, and relative activity per unit protein was about 28 μmol/h. This active fraction contained only protein indicated by a single band of about 90 kDa to 100 kDa molecular weight measured by electrophoresis as described below. The resulting refined enzyme sample had about 2000 times as much relative activity as the raw extract. The yield of enzyme after strongly basic anion exchange chromatography by HiTrapQ was 12%. Table 1 shows the results of refining.

[Table 1]

Table 1

	Total Protein mg	Total Activity nmol/h	Relative Activity nmol/h/mg	Yield %
raw extract	1915	20700	11	—
HiTrapQ	1092	48800	45	100
DEAE-TOYOPEARL	510	33000	61	68
Superdex 200pg	1.79	26500	14800	54
hydroxyapatite chromatography (1)	0.51	12600	24700	26
hydroxyapatite chromatography (2)	0.20	5700	28500	12

[Working Example 2] Study of Characteristics of Raffinose Synthetase

The characteristics of the refined raffinose synthetase obtained in Working Example 1 were studied.

<1> Molecular Weight Measurement

(1) Gel Filtration Chromatography

Taking 10 μL of refined raffinose synthetase, this sample and a molecular weight marker (Gel Filtration Molecular Weight Marker Kit: manufactured by Pharmacia) were supplied to a gel filtration chromatogram (Superdex 200pg; manufactured by Pharmacia). The column was balanced and

eluted using Buffer 6 (20 mM tris hydrochloric acid buffer (pH 7.0), 0.1 M NaCl, 5 mM DTT, 0.02% Tween 20). As a result, the molecular weight of the raffinose synthetase was surmised to be about 75 kDa to 95 kDa.

(2) Polyacrylamide Gel Electrophoresis (Native PAGE)

Ten µL of refined raffinose synthetase was taken and combined with an equal weight of a sample buffer (0.0625 M tris-hydrochloric acid (pH 6.8), 15% glycerol, 0.001% BPB) to form an electrophoresis sample. Ten µL of this sample were supplied to a 10% polyacrylamide gel (manufactured by Daiichi Pure Chemicals, Multigel 10) and electrophoresed in a 0.025 M tris-0.192 M glycine buffer (pH 8.4) at 40 mA for about sixty minutes. After electrophoresing, the sample was stained using a Silvest [as transliterated] Stain Kit (manufactured by Nacalai Tesque). As a result, the molecular weight was surmised to be about 90 kDa to 100 kDa.

(3) SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Ten µL of refined raffinose synthetase were taken and combined with an equal weight of a sample buffer (0.0625 M tris-hydrochloric acid (pH 6.8), 2% SDS, 10% glycerol, 5% mercaptoethanol, 0.001% BPB) and heated in a boiling solution for one minute to form an electrophoresis sample. Ten µL of this sample were supplied to a 10% to 20% gradient polyacrylamide gel (manufactured by Daiichi Pure Chemicals) and electrophoresed in a 0.025 M tris-0.192 M glycine buffer containing 0.1% SDS for about seventy minutes. After electrophoresing, the sample was stained using a /1
Silvest Stain Kit (manufactured by Nacalai Tesque). As a result, the molecular weight was surmised to be about 90 kDa to 100 kDa

<2> Optimum Reaction Temperature

Raffinose synthetase activity was assayed under various temperature conditions (28°C, 32°C, 36°C, 40°C, 48°C, and 52°C) following the method for assaying raffinose synthetase activity described above. Enzyme solution was added to each reaction solution to bring to 2 µL. Fig. 3 shows the relative activity at each temperature taking the enzymatic activity at 32°C as 100. As a result, raffinose synthetase showed activity over a range of about 25°C to 42°C. The optimum reaction temperature was close to 35°C to 40°C.

<3> Optimum Reaction pH

Raffinose synthetase activity was assayed under various pH conditions (pH 4 to 11) following the method for assaying raffinose synthetase activity described above. Fifty mM ascorbic acid buffer (pH 4 to 6), 50 mM potassium phosphate buffer (pH 5.5 to 7.5), 50 mM bis-tris hydrochloric acid buffer (pH 6 to 7), 20 mM tris-hydrochloric acid buffer (pH 7 to 8.5), and 50 mM glycine-NaOH buffer (pH 9 to 11) were used in each reaction. Enzyme solution was added to each reaction solution to bring to 2 µL. The results are given in Fig. 4.

As a result, raffinose synthetase showed activity in a range of pH 5 to 10. The optimum reaction pH differed depending on the type of buffer used, but was close to pH 6 to 8.

<4> Study of Inhibitors and Metallic Ions

Various oxygen inhibitors or metallic ions were added to reaction solutions of refined raffinose synthetase to bring to a final concentration of 1 mM, then raffinose synthetase activity was assayed in the same way as described above. Table 2 shows the residual activity compared to the enzymatic activity when no inhibitor or metallic ions were added. Iodoacetamide markedly inhibited enzymatic activity, and N-ethylmaleimide also showed an inhibitory effect. $CaCl_2$, $CuCl_2$, and $MgCl_2$ showed hardly any inhibitory effect, but $MnCl_2$, $ZnCl_2$, and $NiCl_2$ showed an inhibitory effect.
[Table 2]

Table 2

Inhibitor or Metallic Ion	Residual Activity (%)
none added	100
iodoacetamide	0
N-ethylmaleimide	40
$CaCl_2$	115
$CuCl_2$	101
$MgCl_2$	96
$MnCl_2$	32
$ZnCl_2$	42
$NiCl_2$	68

<5> Inhibition by Myoinositol

Inhibition by myoinositol, which is a reaction product of raffinose synthesis, was investigated. Various concentrations of myoinositol were added to reaction solutions, then raffinose synthetase activity was assayed. Fig. 5 shows the results. Enzymatic activity was inhibited as higher concentrations of myoinositol were added.

<6> Stable pH

The raffinose synthetase fractions obtained in anion exchange chromatography (2) described above were incubated for four hours at 4°C in 50 mM bis-tris hydrochloric acid buffer (pH 5 to 8, containing 0.5 mM DTT) or 20 mM tris-hydrochloric acid buffer (pH 7 to 8, containing 5 mM DTT), then assayed for raffinose synthetase activity. Fig. 6 shows enzymatic activity plotted against the pH of the buffer used for incubating. Raffinose synthetase activity was found under all incubation conditions, and was especially stable in a range of pH 5 to 7.5.

<7> Stable Temperature

The raffinose synthetase fractions obtained in anion exchange chromatography (2) described above were incubated in 20 mM tris-hydrochloric acid buffer (pH 7, containing 5 mM DTT) for sixty minutes at 28°C, 32°C, 37°C, or 40°C, then assayed for raffinose synthetase activity. As a result, this enzyme had 80% to 100% activity compared to a control fraction which was not incubated as described above, and was stable in a range of 28°C to 40°C.

<8> Analysis of Amino Acid Sequence

The cysteine residual group of refined raffinose synthetase was ethylated with reductive pyridyl-ethyl and desalted. The resulting product was consumed by lysyl endopeptidase (Achromobacter protease 1, manufactured by Wako Pure Chemicals) at 37°C for twelve hours to produce peptide fragments. The resulting peptide solution mixture was supplied to reverse-phase HPLC (column: Waters μBondasphere (φ 2.1 x 150 mm, C_{18}, 300), manufactured by Waters (Millipore)) to isolate and acquire peptide fragments. The peptide fragments were eluted by an acetonitrile concentration gradient using 0.1% TFA (trifluoroacetic acid) as the /1
solvent. The amino acid sequence for three of the peptide fragments acquired was determined by a protein sequencer. Sequence numbers 1 to 3 in the sequence charts express the amino acid sequences of these three peptides. These peptides are referred to below as Peptides 1, 2, and 3 in this order.

[Working Example 3] Acquisition of DNA Coding Cucumber-Derived Raffinose Synthetase

<1> Isolation of Partial Fragment of Raffinose Synthetase cDNA by PCR

Twenty-two grams of cucumber leaf vein parts were crushed in liquid nitrogen using a mortar. This crushed product was combined with a mixture of an extraction buffer (100 mM lithium chloride, 100 mM tris-hydrochloric acid (pH 8.0), 10 mM EDTA, 1% SDS) and an equal weight of phenol preheated to 80°C. These were agitated, combined with phenol and an equal eight of chloroform-isoamyl alcohol (24:1), and agitated again. This mixture was centrifuged at 4°C and 9250 x g for fifteen minutes, and the supernatant was collected. This supernatant was treated with phenol and chloroform-isoamyl alcohol again, and the centrifuged supernatant was collected. This supernatant was combined with an equal weight of 4 M lithium chloride and left at -70°C for one hour.

After thawing to room temperature, the mixture was centrifuged at 4°C and 9250 x g for thirty minutes to produce a sediment. This sediment was washed once with 2 M lithium chloride and once with 80% ethanol, dried, then dissolved in 2 mL water treated with diethyl pyrocarbonate and taken as the total refined RNA. The total RNA obtained was 2.38 mg.

The total weight of this total RNA was supplied to a poly(A)$^+$ RNA refining kit using an oligo(dT) cellulose column (manufactured by STRATAGENE CLONING SYSTEMS), and the poly(A)$^+$ RNA molecules were refined to give 42.5 μg of poly(A)$^+$ RNA.

A single-stranded cDNA was synthesized from the poly(A)$^+$ RNA obtained as described above, using the reverse transcription enzyme SuperScriptII (manufactured by GIBCO BRL). This cDNA was amplified by PCR to isolate the raffinose synthetase cDNA from this cDNA mixture. For the primers used in this PCR, the single-stranded oligonucleotides (sequence numbers 6 to 22) shown in Fig. 7 were synthesized from the peptide fragments of cucumber-derived raffinose synthetase determined in Working Example 2. In the sequences of these primers, R represents A or G, Y represents C or T, M represents A or C, K represents G or T, D represents G, A, or T, H represents A, T, or C, B represents T or C, and N represents G, A, G, C, or inosine (base: hypoxanthine).

Recombining A (A1 (sequence number 6), A2 (sequence number 7), A3 (sequence number 8), or A4 (sequence number 9)) with the 5′ side of the primer and D′ (D′1 (sequence number 21) or D′2 (sequence number 22)) with the 3′ side of the primer as a primer amplified a DNA of about 540 base pairs when C2 (sequence number 14) was used on the 5′ side of the primer and B′1 (sequence number 18) or B′2 (sequence number 19) was used on the 3′ side of the primer. Cloning this fragment to plasmid pCRII using a TA cloning kit (manufactured by INVITROGEN BV) and analyzing the base sequences revealed base sequences coding the amino acid sequences of peptides 1 and 2 just before the primer sequences of the two terminals, indicating that the above-mentioned amplified fragment was a DNA fragment derived from the gene for raffinose synthetase.

Furthermore, 3′ RACE was performed using a RACE kit (3′ Ampifinder RACE Kit (manufactured by CLONTECH)) to determine the position of the above-mentioned cloned PCR-amplified DNA fragment on the gene for raffinose synthetase.

PCR was performed with the above-mentioned cDNA mixture as a matrix, using a primer having C (C1 (sequence number 13) or C2 (sequence number 14)) on the 5′ side of the primer and oligo(dT) and an anchor sequence on the 3′ side of the primer. Next, PCR was performed with the amplified fragment obtained in this way as a matrix, using a primer having D (D1 (sequence number 15) or D2 (sequence number 16)), which is located further inside than C, on the 5′ side of the primer and oligo(dT) and an anchor sequence on the 3′ side of the primer. As a result, a DNA fragment of about 24000 base pairs was amplified only when PCR was performed by a D2 (sequence number 16) and oligo(dT)anchor primer with DNA amplified by a C1 (sequence number 13) or C2 (sequence number 14) and oligo(dT)-anchor primer as a matrix. PCR was also performed using a primer having C (C1 (sequence number 13) or C2 (sequence number 14)) on the 5′ side of the primer and oligo(dT) and an anchor sequence on the 3′ side of the primer,

then PCR was performed with the amplified fragment obtained in this way as a matrix, using a primer having E (sequence number 17) on the 5′ side of the primer and oligo(dT) and an anchor sequence on the 3′ side of the primer. As a result, a DNA fragment of about 300 base pairs was amplified in both cases.

Similarly, PCR was performed with the above-mentioned cDNA mixture as a matrix, using a primer having A (A1 (sequence number 6), A2 (sequence number 7), A3 (sequence number 8), or A4 (sequence number 9)) on the 5′ side of the primer and oligo(dT) and an anchor sequence on the 3′ side of the primer. Next, PCR was performed with the amplified fragment obtained in this way as a matrix, using a primer having B (B1 (sequence number 10), B2 (sequence number 11), or B3 (sequence number 12)), which is located further inside than A, on the 5′ side of the primer and the same oligo(dT) and an anchor sequence on the 3′ side of the primer. This cloned DNA fragments amplified using A2 and B2 primers. Analysis of the base /1
sequences revealed a base sequence coding the amino acid sequence of peptide fragment 1 used to prepare the primer on the 5′ side, and also revealed a base sequence corresponding to peptide fragment 3 above this.

Combined with the earlier PCR results, this shows that the raffinose synthetase peptide fragments are arranged in the order of 2, 1, and 3 from the N-terminal side, and are located near the 5′ terminal on the gene for raffinose synthetase. Since it is desirable that the DNA used as a probe to screen for a cDNA clone containing the total length of the gene for raffinose synthetase be able to detect the part close to 5′ terminal, this DNA fragment was used as a probe to screen a cDNA library.

<2> Cloning Total Length of Region Coding Raffinose Synthetase cDNA

First, a cDNA library was prepared as follows: A double-stranded cDNA was synthesized from 3.8 μg of the poly(A)$^+$ RNA obtained in <1> using a TimeSaver cDNA synthesis kit (manufactured by Pharmacia Biotech). The resulting cDNA was recombined with the EcoRI restriction enzyme cutting site of the λ phage vector λMOSSlox (manufactured by Amersham), then taken up into the phage protein using a Gigapack II Gold packaging kit (manufactured by STRATAGENE CLONING SYSTEMS) to prepare a cucumber cDNA library. Moreover, the titer of this library was 1.46×10^7 pfu/Mg vector.

The host cell Escherichia coli ER1647 was infected with a phage corresponding to 4×10^5 pfu from the above-mentioned cucumber cDNA library, then sown in fourteen 90-mm diameter agar plates at 1.0×10^4 pfu per plate. After culturing at 37°C for about 6.5 hours, the phage plaques formed on these plates were transferred to Nylon membranes (Hybond-N$^+$ manufactured by Amersham).

Next, the above-mentioned Nylon membranes were treated with alkali to denature the transcribed DNA, neutralized, then washed. Following this, the DNA was fixed on top of these Nylon membranes by treating for two

hours at 80°C.

The resulting Nylon membranes were screened for anionic clones using the DNA fragment of about 540 base pairs obtained in <1> as a probe. The DNA fragment of about 540 base pairs described above was consumed by restriction enzyme EcoRI, then electrophoresed. Just the insert of about 540 base pairs was cut out and refined, then labeled with fluorescein using a DNA labeling and detection system (Gene Images Labeling and Detection System (manufactured by Amersham)) to form a probe. The above-mentioned Nylon membranes were pre-hybridized at 60°C for thirty minutes, then combined with the labeled probe and hybridized at 60°C for sixteen hours. An antibody for detecting labeled DNA (alkali phosphatase-labeled anti-fluorescein antibody) was used after diluting 50,000 times. This screening yielding several anionic clone candidates. The resulting candidates were screened twice again as described above to acquire purified anionic clones.

Escherichia coli BM25.8 was infected with the above-mentioned anionic clones, and plasmid vector pMOSSlox-CRS was extracted by culturing on a selection medium containing carbenicillin. The insert cDNA of this plasmid had a length of about 2.5 kb. Next, Escherichia coli MJ109 was transformed by this plasmid. A plasmid DNA prepared from this transformant formed a sample for analyzing the base sequence.

The base sequence of the insert cDNA was analyzed by a conventional method using a Taq Dye Deoxy Terminator Cycle Sequencing Kit (manufactured by Perkin-Elmer).

As a result, this revealed a base sequence comprised of 2352 base pairs expressed by sequence number 4 in the sequence charts. This sequence was found to contain a part matching the base sequence of the DNA probed used by the present inventors. Sequence numbers 4 and 5 show amino acid sequences translated from the base sequence. This amino acid sequence was found to contain parts matching Peptide 1 (amino acid sequence 215 to 244 in sequence number 5), Peptide 2(amino acid sequence 61 to 79 in sequence number 5), and Peptide 3 (amino acid sequence 756 to 769 of sequence number 5) of the cucumber-derived raffinose synthetase obtained by the present inventors, confirming that these peptides code raffinose synthetase.

Transformant AJ13263 of Escherichia coli JM109 supporting plasmid pMossloxCRS containing a DNA coding raffinose synthetase obtained as described above was deposited for international safekeeping based on the Budapest Convention with the National Institute of Bioscience and Human Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry (Post Office Box 305, 1-1-3 Higashi, Tsukuba-shi, Ibaraki Prefecture, Japan) on November 19, 1996, and assigned the receipt number of FERM BP-5748.

[Working Example 4] Acquisition of DNA Coding Soybean-Derived Raffinose Synthetase

<1> Searching a Probe for Cloning Gene for Soybean-Derived Raffinose Synthetase

The total RNA of soybean was Northern-hybridized using the total length of the gene for cucumber-derived raffinose synthetase as a probe. For the probe, a DNA fragment containing the gene for cucumber- /1
derived raffinose synthetase was obtained by consuming plasmid pMossloxCRS obtained in Working Example 3 by restriction enzyme NotI, subjecting the result to agarose gel electrophoresis, isolating the insert fragment, and labeling the isolated DNA fragment by α-P^{32} dCTP. The total RNA was 30 µg of total RNA prepared by the SDS-phenol method from immature soybean seeds (five to six weeks after blossoming). The total RNA was pre-hybridized for thirty minutes, then combined with the probe and hybridized at 42°C overnight. The result was washed under conditions of 1 x SSC, 0.1% SDS, and 60°C. The signal of the control cucumber-derived raffinose synthetase was found, but no clear signal for soybean-derived RNA was obtained. It was decided that probing a better preserved region would be better than probing the total length of cucumber.

<2> Isolation of Partial Fragment of the Gene for Raffinose Synthetase in Arabidopsis

125 mg Arabidopsis pods 17 to 20 days after blossoming were crushed in liquid nitrogen using a mortar. The crushed product was combined with 3 mL of 2 x CTAB (2% cetyl trimethyl ammonium bromide, 0.1 M tris hydrochloric acid buffer (pH 9.5), 1.4 M NaCl, 0.5% mercaptoethanol) and dispersed, then shaken at 65°C for ten minutes. This mixture was transferred to a (50 mL) blue max Falcon tube, combined with 3 mL chloroform-isoamyl alcohol (24:1 (v/v)) and gently mixed, then centrifuged at 12,000 rpm for ten minutes. The supernatant was extracted again by 3 mL chloroform-isoamyl alcohol (24:1 (v/v)) and centrifuged at 10,000 rpm for twenty-five minutes. 1.5 mL Isoamyl alcohol were combined and mixed with 1.8 mL of the centrifuged supernatant, then centrifuged at 12,000 rpm for fifteen minutes at 4°C to give a sediment. The sediment was washed with 70% ethanol and dried, then dissolved in 1 mL of TE buffer. This solution was combined and mixed with one-fourth its weight of 10 M lithium chloride solution, left in ice for four hours, then centrifuged at 12,000 rpm for fifteen minutes at 4°C. The sediment was washed with 2 M lithium chloride and 70% ethanol and dried, then dissolved in 100 µL of TE buffer. This solution was combined with phenol-chloroform (1:1 (v/v)) and agitated, then centrifuged at 12,000 rpm for fifteen minutes at 4°C to obtain an aqueous layer. This aqueous layer was applied to the ethanol sediment, and the sediment was washed with 70% ethanol, then dried and dissolved in 10 µL of water treated with diethyl pyrocarbonate to give the total RNA. The total RNA obtained was 18.7 µg. A single-stranded DNA was synthesized from this total RNA using the reverse transcription

enzyme SuperScriptII (manufactured by GIBCO BRL).

A primer was synthesized for amplifying a partial fragment of the gene for raffinose synthetase from this cDNA mixture by PCR. The primer was synthesized as a single-stranded oligonucleotide (sequence numbers 25 and 26) from the preserved region by searching the gene bank for a DNA which is homologous with the gene for cucumber-derived raffinose synthetase. Performing PCR by this primer with the previous single-stranded cDNA as a matrix amplified a DNA fragment of about 250 base pairs. Cloning this fragment to plasmid pCR2.1 using a TA cloning kit (manufactured by INVITROGEN BV) and analyzing the base sequence gave the base sequence expressed by sequence number 27. This indicated that the product was a partial fragment of Arabidopsis-derived raffinose synthetase cDNA which is homologous with cucumber-derived raffinose synthetase.

<3> Cloning Soybean-Derived Raffinose Synthetase cDNA

4.5 g Soybean seeds five to six weeks after blossoming were crushed in liquid nitrogen using a mortar. 1.3 mg Total RNA were extracted from the crushed product by the SDS-phenol method. The extract was supplied to a poly(A)$^+$ RNA refining kit using an oligo(dT) cellulose column (manufactured by STRATAGENE CLONING SYSTEMS), and about 6 μg of poly(A)$^+$ RNA were isolated. A double-stranded cDNA was synthesized from about 2 μg of this poly(A)$^+$ RNA by an oligo dT primer using a TimeSaver cDNA synthesis kit (manufactured by Pharmacia Biotech). The resulting cDNA was recombined with the EcoRI restriction enzyme cutting site of the λ phage vector λMOSSlox (manufactured by Amersham), then taken up into the phage grains using a Gigapack III Gold packaging kit (manufactured by STRATAGENE CLONING SYSTEMS) to prepare a soybean cDNA library. Moreover, the titer of this library was 1.42×10^7 pfu/Mg vector DNA.

A phage corresponding to 1.5×10^5 pfu from the soybean cDNA library was transferred and fixed to Nylon membranes (Hybond-N$^+$ manufactured by Amersham) in the same way as the cucumber cDNA library. Two membranes were transferred per plate, and two sets were prepared. The resulting membranes were screened using the partial gene fragment of Arabidopsis-derived raffinose synthetase cDNA obtained in <2>. This DNA fragment was labeled with fluorescein using a Gene Images Labeling and Detection System (manufactured by Amersham) to form a probe. The labeled DNA fragments were hybridized and detected in the same way as when screening the cucumber-derived cDNA library. However, one set of the membranes was washed at 60°C by 1 x SSC and 0.1% SDS, and the other set was washed at 60°C by 0.1 x SSC and 0.1% SDS. Each conditions produced fifteen cloning candidates. The candidates were screened once more in the same way as described above to acquire five purified clone strains.

E. coli BM25.8 was infected with the above-mentioned anionic clones and the plasmid containing cDNA was cut out. Furthermore, E. coli JM109 was transformed by this plasmid, and plasmid DNA was prepared from /2

the transformant to form a sample. The base sequence was analyzed in the same way as the gene for cucumber-derived raffinose synthetase. Of the five clones, one, pMOSSloxSRS, was found to contain the total length of the gene for soybean-derived raffinose synthetase.

The insert fragment of pMOSSloxSRS has a sequence comprised of 2780 base pairs expressed by sequence number 23, and codes a raffinose synthetase comprised of 750 amino acids. The insert fragments of the other clones were shorter than pMOSSloxSRS, and lacked the 5' side of the gene for raffinose synthetase.

Transformant AJ13379 of Escherichia coli JM109 supporting plasmid pMossloxSRS containing a DNA coding raffinose synthetase obtained as described above was deposited for international safekeeping based on the Budapest Convention with the National Institute of Bioscience and Human Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry (Post Office Box 305, 1-1-3 Higashi, Tsukuba-shi, Ibaraki Prefecture, Japan) on November 19, 1996, and assigned the receipt number of FERM BP-6149.

[Working Example 5] Chimera Gene Containing DNA Coding Raffinose Synthetase, and Transformed Plant

<1> Construction of Plasmid Containing Chimera Gene

A DNA fragment of Arabidopsis-derived raffinose synthetase was introduced into the Agrobacterium LBA4404 using pBI121 (CLONTECH) as a binary vector. pBI121 is a plasmid derived from pBIN19 connecting a nopalin synthetase promoter, a gene with a neomycin phosphotransferase configuration (NPTII), the terminator of the gene for nopalin synthetase (Nos-ter), a CaMV 35S promoter, the gene for GUS (β-glucuronidase), and Nos-ter, and has sequences on both terminals which can be transferred to a plant. This plasmid has a SmaI site below the CaMV 35S promoter, and an insert introduced at this site manifests the gene under the control of said promoter.

The fragment of the gene for cucumber-derived raffinose synthetase obtained in Working Example 3 was introduced into binary vector pBI121. The gene for raffinose synthetase was consumed by DraI, and a DNA fragment containing bases from the 1382^{nd} base to the 2529^{th} base in sequence number 4 in the sequence charts was prepared by agarose gel electrophoresis. Escherichia coli HB101 was transformed using this ligation reaction solution, and recombinant plasmids were prepared from the transformed strain. Of the resulting recombinant plasmids, two types were selected, one in which the fragment of raffinose synthetase DNA was connected to the CaMV 35S promoter in reverse order (antisense) and one in which this DNA was connected in the regular order (sense). These were named pBIcRS1 and pBIcRS9, respectively.

A plasmid containing a chimera gene for manifesting raffinose

synthetase was also constructed. pMOSSloxCRS containing the gene for raffinose synthetase was consumed by NotI, and a fragment of the gene for raffinose synthetase was prepared by agarose gel electrophoresis. The NotI cutting site on this DNA fragment was filled in by a Taq polymerase reaction using dNTP, producing a CRS fragment with a projecting A base on the 3′ side. Separately, pBI121 was consumed by SmaI, a straight-chain DNA was refined by agarose gel electrophoresis, and pBI121/SmaI with a T base appended to 3′ side was obtained by a Taq polymerase reaction using dTTP. After refining, the CRS fragment was ligated to pBI121/SmaI. Escherichia coli HB101 was transformed using this ligation reaction solution. A plasmid DNA was prepared from the resulting transformant, and this plasmid DNA was consumed by each of restriction enzymes EcoRI, BamHI, and XhoI and by recombining. The molecular weight of the resulting fragment was measured by agarose electrophoresis, and a physical map was prepared. Of the recombinant plasmids, the recombinant plasmid for the CaMV 35S promoter, which has the gene for raffinose synthetase connected in the regular order, was selected from the prepared physical map, and was named pBIsRS1.

The plasmids obtained as described above were introduced into Agrobacterium LBA4404 by triparental mating of Escherichia HB101 and Agrobacterium LBA4404 containing these.

<2> Transformation

Arabidopsis was transformed as follows: Seeds of Arabidopsis thaliana were hydrated, then disinfected by treating for five minutes in 80% ethanol containing 1% Tween 20 and for ten minutes in 10% sodium chlorite likewise containing 1% Tween 20, then washing five times with sterilized water. The seeds were suspended in 1% low-melting-point agarose, then sown in an MS medium (basic MS medium (Murashige and Skoog, *Physiologia Plantrum*, 15, 473-497 (1962)), Vitamin B5, 10 g/L sucrose, 0.5 g/L MES, pH 5.8). After culturing at 22°C for one week in a culture room which was exposed to light for sixteen hours and kept dark for eight hours a day, the seeds that had sprouted were transplanted into rock fiber. Culture of the sprouted seeds was continued under the same conditions. After three weeks, the plants had pushed up, and the hearts were picked when the stem height reached several centimeters. After picking the hearts, the plants were grown for one week until the first flowers on the side branches had blossomed.

Agrobacterium with recombinant plasmid containing the gene for raffinose synthetase introduced was pre-cultured in 2 mL LB medium. /2
This was inoculated with an LB culture containing 50 mg/L kanamycin and 25 mg/L streptomycin, then cultured at 28°C for about one day. Bacteria were collected at room temperature, and the biomass was suspended in a suspension medium for infiltration (1/2 MS salt, 1/2 Gamborg Vitamin B5, 5% sucrose, 0.5 g/L MES, pH 5.7 (KOH); benzyl aminopurine added to a final

concentration of 0.044 µM and Silwet L77 added at 200 µL per liter (final concentration: 0.02%) before use) until OD_{600} was 0.8.

Blossoming and fruiting flowers were picked from the plant to be infiltrated. The rock fiber was inverted and non-fruiting flowers were soaked in the above-mentioned Agrobacterium suspension. A desiccator was added and the suspension was reduced in pressure to 40 mmHG for fifteen minutes. Seeds were collected for two to four weeks. The harvested seeds were stored in a desiccator.

Next, transformants were selected in the suspension medium. The seeds described above were sterilized, then cultured at 22°C in a selection medium (MS salt, Gamborg Vitamin B5, 1% sucrose, 0.5 g/L MES, pH 5.8, 0.8% agar; selection antibiotics, carbenicillin (final concentration: 100 mg/L) and kanamycin (final concentration: 50 mg/L), added after autoclaving). Resistant plants were selected. The resistant plants were transferred to a fresh medium and grown until they developed main leaves. Seeds were harvested from these plants. Selection was repeated in the same way, and T3 seeds were harvested. The T3 seeds were assayed for raffinose content in the same way as described above. Table 3 shows the results.

[Table 3]

Table 3

Plant (plasmid)	Raffinose Content (mg/g)
wild strain	0.20
transformant (pBIcRS1)	0.00
transformant (pBIcRS9)	0.00
transformant (pBIsRS1)	0.22

[Effects of the Invention]

The present invention offers refined raffinose synthetase, a gene for raffinose synthetase, a chimera gene having a gene for raffinose synthetase and a transcription control region which can be manifested in plant cells, and a plant into which this chimera gene has been introduced.

The raffinose synthetase of the present invention can be used to efficiently synthesize raffinose from sucrose and galactinol. The gene for raffinose synthetase or the chimera gene of the present invention can also be used to vary the raffinose family oligosaccharide content of a plant.

[Sequence Charts]

Sequence Number: 1
Length of Sequence: 30

Type of Sequence: amino acid
Typology: straight chain
Category of Sequence: peptide
Fragment Type: intermediate peptide
Sequence

Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr Leu Thr Val His Pro Gln
1 5 10 15
Gly Val Ile Glu Gly Val Arg His Leu Val Asp Gly Gly Cys
20 25 30

Sequence Number: 2
Length of Sequence: 19
Type of Sequence: amino acid
Typology: straight chain
Category of Sequence: peptide
Fragment Type: intermediate peptide
Sequence

Pro Val Ser Val Gly Cys Phe Val Gly Phe Asp Ala Ser Glu Pro Asp
1 5 10 15
Ser Arg His

Sequence Number: 3
Length of Sequence: 14
Type of Sequence: amino acid
Typology: straight chain
Category of Sequence: peptide
Fragment Type: intermediate peptide
Sequence /2

Tyr Asp Gln Asp Gln Met Val Val Val Gln Val Pro Trp Pro
1 5 10

Sequence Number: 4
Length of Sequence: 2569
Type of Sequence: nucleic acid
Typology: double strand
Category of Sequence: cDNA to mRNA
Origin
Name of Organism: cucumber (Cucumis sativus)
Characteristics of Sequence
Symbol Expressing Characteristics: CDS
Positions Where Located: 56 … 2407

Sequence

```
AAAAAACAAC CCTTCTTTTA GTTTTTTGGG TTTGTTTCTT CTTTTCTTCT CACAA ATG    58
                                                             Met
                                                               1
GCT CCT AGT TTT AAA AAT GGT GGC TCC AAC GTA GTT TCA TTT GAT GGC    106
Ala Pro Ser Phe Lys Asn Gly Gly Ser Asn Val Val Ser Phe Asp Gly
            5                  10                  15
TTA AAT GAC ATG TCG TCA CCG TTT GCA ATC GAC GGA TCG GAT TTC ACT    154
Leu Asn Asp Met Ser Ser Pro Phe Ala Ile Asp Gly Ser Asp Phe Thr
        20                  25                  30
GTG AAC GGT CAT TCG TTT CTG TCC GAT GTT CCT GAG AAC ATT GTT GCT    202
Val Asn Gly His Ser Phe Leu Ser Asp Val Pro Glu Asn Ile Val Ala
    35                  40                  45
TCT CCT TCT CCG TAC ACT TCG ATA GAC AAG TCC CCG GTT TCG GTT GGT    250
Ser Pro Ser Pro Tyr Thr Ser Ile Asp Lys Ser Pro Val Ser Val Gly
 50                  55                  60                  65
TGC TTT GTT GGA TTC GAC GCG TCG GAA CCT GAT AGC CGA CAT GTT GTT    298
Cys Phe Val Gly Phe Asp Ala Ser Glu Pro Asp Ser Arg His Val Val
                70                  75                  80
TCG ATT GGG AAG CTG AAG GAT ATT CGG TTT ATG AGT ATT TTC AGG TTT    346
Ser Ile Gly Lys Leu Lys Asp Ile Arg Phe Met Ser Ile Phe Arg Phe
            85                  90                  95
AAG GTT TGG TGG ACT ACA CAC TGG GTT GGT CGA AAT GGT GGG GAT CTT    394
Lys Val Trp Trp Thr Thr His Trp Val Gly Arg Asn Gly Gly Asp Leu
        100                 105                 110
GAA TCG GAG ACT CAG ATT GTG ATC CTT GAG AAG TCA GAT TCT GGT CGA    442
Glu Ser Glu Thr Gln Ile Val Ile Leu Glu Lys Ser Asp Ser Gly Arg
    115                 120                 125
CCG TAT GTT TTC CTT CTT CCG ATC GTT GAG GGA CCG TTC CGA ACC TCG    490
Pro Tyr Val Phe Leu Leu Pro Ile Val Glu Gly Pro Phe Arg Thr Ser
130                 135                 140                 145
ATT CAG CCT GGG GAT GAT GAC TTT GTC GAT GTT TGT GTC GAG AGT GGT    538
Ile Gln Pro Gly Asp Asp Asp Phe Val Asp Val Cys Val Glu Ser Gly
                150                 155                 160
TCG TCG AAA GTT GTT GAT GCA TCG TTC CGA AGT ATG TTG TAT CTT CAT    586
Ser Ser Lys Val Val Asp Ala Ser Phe Arg Ser Met Leu Tyr Leu His
            165                 170                 175
GCT GGT GAT GAT CCG TTT GCA CTT GTT AAA GAG GCG ATG AAG ATC GTG    634
Ala Gly Asp Asp Pro Phe Ala Leu Val Lys Glu Ala Met Lys Ile Val
        180                 185                 190
AGG ACC CAT CTT GGA ACT TTT CGC TTG TTG GAG GAG AAG ACT CCA CCA    682
```

```
Arg Thr His Leu Gly Thr Phe Arg Leu Leu Glu Glu Lys Thr Pro Pro
    195                 200                 205
GGT ATC GTG GAC AAA TTC GGT TGG TGC ACG TGG GAC GCG TTT TAC CTA     730
Gly Ile Val Asp Lys Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr Leu
210                 215                 220                 225
ACG GTT CAT CCA CAG GGC GTA ATA GAA GGC GTG AGG CAT CTC GTC GAC     778
Thr Val His Pro Gln Gly Val Ile Glu Gly Val Arg His Leu Val Asp
                230                 235                 240
GGC GGT TGT CCT CCC GGT TTA GTC CTA ATC GAC GAT GGT TGG CAA TCC     826
Gly Gly Cys Pro Pro Gly Leu Val Leu Ile Asp Asp Gly Trp Gln Ser
            245                 250                 255
ATC GGA CAC GAT TCG GAT CCC ATC ACC AAA GAA GGA ATG AAC CAA ACC     874
Ile Gly His Asp Ser Asp Pro Ile Thr Lys Glu Gly Met Asn Gln Thr
        260                 265                 270
GTC GCC GGC GAG CAA ATG CCC TGC CGT CTT TTG AAA TTC CAA GAG AAT     922
Val Ala Gly Glu Gln Met Pro Cys Arg Leu Leu Lys Phe Gln Glu Asn
    275                 280                 285
TAC AAA TTC CGT GAC TAC GTC AAT CCC AAG GCC ACC GGC CCC CGA GCC     970
Tyr Lys Phe Arg Asp Tyr Val Asn Pro Lys Ala Thr Gly Pro Arg Ala
290                 295                 300                 305
GGC CAG AAG GGG ATG AAG GCG TTT ATA GAT GAA CTC AAA GGA GAG TTT    1018
Gly Gln Lys Gly Met Lys Ala Phe Ile Asp Glu Leu Lys Gly Glu Phe
                310                 315                 320
AAG ACT GTG GAG CAT GTT TAT GTT TGG CAT GCT TTG TGT GGA TAT TGG    1066
Lys Thr Val Glu His Val Tyr Val Trp His Ala Leu Cys Gly Tyr Trp
            325                 330                 335
GGT GGC CTT CGC CCG CAG GTG CCT GGC TTG CCT GAG GCA CGT GTG ATT    1114
Gly Gly Leu Arg Pro Gln Val Pro Gly Leu Pro Glu Ala Arg Val Ile
        340                 345                 350
CAG CCA GTG CTT TCA CCA GGG CTG CAG ATG ACG ATG GAG GAT TTG GCG    1162
Gln Pro Val Leu Ser Pro Gly Leu Gln Met Thr Met Glu Asp Leu Ala
    355                 360                 365
GTG GAT AAG ATT GTT CTT CAT AAG GTC GGG CTG GTC CCG CCG GAG AAG    1210
Val Asp Lys Ile Val Leu His Lys Val Gly Leu Val Pro Pro Glu Lys
370                 375                 380                 385
GCT GAG GAG ATG TAC GAA GGA CTT CAT GCT CAT TTG GAA AAA GTT GGG    1258
Ala Glu Glu Met Tyr Glu Gly Leu His Ala His Leu Glu Lys Val Gly
                390                 395                 400
ATC GAC GGT GTT AAG ATT GAC GTT ATC CAC CTA TTG GAG ATG TTG TGT    1306
Ile Asp Gly Val Lys Ile Asp Val Ile His Leu Leu Glu Met Leu Cys
            405                 410                 415
GAA GAC TAT GGA GGG AGA GTG GAT TTG GCA AAG GCA TAT TAC AAA GCA    1354
Glu Asp Tyr Gly Gly Arg Val Asp Leu Ala Lys Ala Tyr Tyr Lys Ala
        420                 425                 430
ATG ACC AAA TCA ATA AAT AAA CAT TTT AAA GGA AAT GGA GTC ATT GCA    1402
Met Thr Lys Ser Ile Asn Lys His Phe Lys Gly Asn Gly Val Ile Ala
    435                 440                 445
AGT ATG GAA CAT TGT AAC GAC TTC ATG TTC CTT GGC ACG GAA GCT ATC    1450
Ser Met Glu His Cys Asn Asp Phe Met Phe Leu Gly Thr Glu Ala Ile
450                 455                 460                 465
```

TCT CTT GGT CGT GTT GGT GAT GAC TTT TGG TGC ACG GAC CCC TCT GGT 1498
Ser Leu Gly Arg Val Gly Asp Asp Phe Trp Cys Thr Asp Pro Ser Gly
470 475 480

GAT CCA AAC GGT ACG TTT TGG CTC CAA GGA TGT CAC ATG GTT CAT TGT 1546
Asp Pro Asn Gly Thr Phe Trp Leu Gln Gly Cys His Met Val His Cys
485 490 495

GCC AAC GAC AGC TTG TGG ATG GGG AAC TTC ATC CAC CCT GAC TGG GAT 1594
Ala Asn Asp Ser Leu Trp Met Gly Asn Phe Ile His Pro Asp Trp Asp
500 505 510

ATG TTC CAA TCC ACC CAC CCT TGT GCC GCC TTC CAT GCT GCC TCT CGA 1642
Met Phe Gln Ser Thr His Pro Cys Ala Ala Phe His Ala Ala Ser Arg
515 520 525

GCC ATC TCT GGT GGC CCG ATC TAT GTT AGT GAT TCT GTG GGA AAG CAT 1690
Ala Ile Ser Gly Gly Pro Ile Tyr Val Ser Asp Ser Val Gly Lys His
530 535 540 545

AAC TTT GAT CTT CTG AAA AAA CTA GTG CTT CCT GAT GGA TCG ATC CTT 1738
Asn Phe Asp Leu Leu Lys Lys Leu Val Leu Pro Asp Gly Ser Ile Leu
550 555 560

CGA AGT GAG TAC TAT GCA CTC CCG ACT CGC GAT TGT TTG TTT GAA GAC 1786
Arg Ser Glu Tyr Tyr Ala Leu Pro Thr Arg Asp Cys Leu Phe Glu Asp
565 570 575

CCT TTG CAT AAT GGA GAA ACT ATG CTT AAG ATT TGG AAT CTC AAC AAG 1834
Pro Leu His Asn Gly Glu Thr Met Leu Lys Ile Trp Asn Leu Asn Lys
580 585 590

TTC ACT GGA GTG ATT GGT GCA TTC AAC TGC CAA GGA GGA GGA TGG TGT 1882
Phe Thr Gly Val Ile Gly Ala Phe Asn Cys Gln Gly Gly Gly Trp Cys
595 600 605

CGT GAG ACA CGC CGC AAC CAA TGC TTT TCA CAA TAC TCA AAA CGA GTG 1930
Arg Glu Thr Arg Arg Asn Gln Cys Phe Ser Gln Tyr Ser Lys Arg Val
610 615 620 625

ACA TCC AAA ACT AAC CCA AAA GAC ATA GAA TGG CAC AGT GGA GAA AAC 1978
Thr Ser Lys Thr Asn Pro Lys Asp Ile Glu Trp His Ser Gly Glu Asn
630 635 640

CCT ATC TCT ATT GAA GGC GTT AAA ACC TTT GCG CTT TAC CTC TAT CAA 2026
Pro Ile Ser Ile Glu Gly Val Lys Thr Phe Ala Leu Tyr Leu Tyr Gln
645 650 655

GCC AAA AAA CTT ATC CTC TCC AAG CCC TCT CAA GAT CTT GAC ATA GCT 2074
Ala Lys Lys Leu Ile Leu Ser Lys Pro Ser Gln Asp Leu Asp Ile Ala
660 665 670

CTT GAC CCA TTC GAA TTC GAG CTC ATC ACT GTT TCA CCA GTG ACC AAA 2122
Leu Asp Pro Phe Glu Phe Glu Leu Ile Thr Val Ser Pro Val Thr Lys
675 680 685

CTC ATC CAA ACT TCT CTA CAC TTT GCC CCA ATT GGG CTG GTG AAC ATG 2170
Leu Ile Gln Thr Ser Leu His Phe Ala Pro Ile Gly Leu Val Asn Met
690 695 700 705

CTT AAC ACT AGT GGA GCC ATC CAA TCT GTG GAC TAT GAC GAT GAC CTA 2218
Leu Asn Thr Ser Gly Ala Ile Gln Ser Val Asp Tyr Asp Asp Asp Leu
710 715 720

AGC TCA GTC GAG ATT GGT GTC AAA GGG TGT GGT GAG ATG CGA GTA TTT 2266
Ser Ser Val Glu Ile Gly Val Lys Gly Cys Gly Glu Met Arg Val Phe

/2

```
        725                 730                 735
GCA TCG AAA AAA CCA AGG GCT TGT CGT ATT GAT GGG GAG GAT GTT GGG    2314
Ala Ser Lys Lys Pro Arg Ala Cys Arg Ile Asp Gly Glu Asp Val Gly
    740                 745                 750
TTC AAG TAT GAT CAG GAC CAA ATG GTG GTG GTT CAA GTG CCA TGG CCA    2362
Phe Lys Tyr Asp Gln Asp Gln Met Val Val Val Gln Val Pro Trp Pro
755                 760                 765
ATT GAT TCT TCA TCG GGT GGC ATT TCG GTT ATC GAG TAC TTG TTT        2407
Ile Asp Ser Ser Ser Gly Gly Ile Ser Val Ile Glu Tyr Leu Phe
770                 775                 780
TAATTTTTAT TTATGTARAG CTCAATGATT GTTGTTGTTG TCGCTGTTGT TGCTATCAAT  2467
GTATTTCTCT CCAAAAGAAA ATTATGTGTA ATTTGGAGAG TAATTAAGTG AGTKAAATTT  2527
TAAATAARAC TACTTTTAAT TATTTATCAA AAAAAAAAAA AA                     2569
```

Sequence Number: 5
Length of Sequence: 784
Type of Sequence: amino acid
Topology: straight chain
Category of Sequence: protein

Sequence

Met Ala Pro Ser Phe Lys Asn Gly Gly Ser Asn Val Val Ser Phe Asp
1 5 10 15
Gly Leu Asn Asp Met Ser Ser Pro Phe Ala Ile Asp Gly Ser Asp Phe
20 25 30
Thr Val Asn Gly His Ser Phe Leu Ser Asp Val Pro Glu Asn Ile Val
35 40 45
Ala Ser Pro Ser Pro Tyr Thr Ser Ile Asp Lys Ser Pro Val Ser Val
50 55 60
Gly Cys Phe Val Gly Phe Asp Ala Ser Glu Pro Asp Ser Arg His Val
65 70 75 80
Val Ser Ile Gly Lys Leu Lys Asp Ile Arg Phe Met Ser Ile Phe Arg
85 90 95
Phe Lys Val Trp Trp Thr Thr His Trp Val Gly Arg Asn Gly Gly Asp
100 105 110
Leu Glu Ser Glu Thr Gln Ile Val Ile Leu Glu Lys Ser Asp Ser Gly
115 120 125
Arg Pro Tyr Val Phe Leu Leu Pro Ile Val Glu Gly Pro Phe Arg Thr
130 135 140
Ser Ile Gln Pro Gly Asp Asp Asp Phe Val Asp Val Cys Val Glu Ser
145 150 155 160
Gly Ser Ser Lys Val Val Asp Ala Ser Phe Arg Ser Met Leu Tyr Leu
165 170 175
His Ala Gly Asp Asp Pro Phe Ala Leu Val Lys Glu Ala Met Lys Ile
180 185 190
Val Arg Thr His Leu Gly Thr Phe Arg Leu Leu Glu Glu Lys Thr Pro
195 200 205
Pro Gly Ile Val Asp Lys Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr
210 215 220
Leu Thr Val His Pro Gln Gly Val Ile Glu Gly Val Arg His Leu Val
225 230 235 240
Asp Gly Gly Cys Pro Pro Gly Leu Val Leu Ile Asp Asp Gly Trp Gln
245 250 255
Ser Ile Gly His Asp Ser Asp Pro Ile Thr Lys Glu Gly Met Asn Gln

260 265 270

Thr Val Ala Gly Glu Gln Met Pro Cys Arg Leu Leu Lys Phe Gln Glu

275 280 285

Asn Tyr Lys Phe Arg Asp Tyr Val Asn Pro Lys Ala Thr Gly Pro Arg

290 295 300

Ala Gly Gln Lys Gly Met Lys Ala Phe Ile Asp Glu Leu Lys Gly Glu

305 310 315 320

Phe Lys Thr Val Glu His Val Tyr Val Trp His Ala Leu Cys Gly Tyr

325 330 335

Trp Gly Gly Leu Arg Pro Gln Val Pro Gly Leu Pro Glu Ala Arg Val

340 345 350

Ile Gln Pro Val Leu Ser Pro Gly Leu Gln Met Thr Met Glu Asp Leu

355 360 365

Ala Val Asp Lys Ile Val Leu His Lys Val Gly Leu Val Pro Pro Glu

370 375 380

Lys Ala Glu Glu Met Tyr Glu Gly Leu His Ala His Leu Glu Lys Val

385 390 395 400

Gly Ile Asp Gly Val Lys Ile Asp Val Ile His Leu Leu Glu Met Leu

405 410 415

Cys Glu Asp Tyr Gly Gly Arg Val Asp Leu Ala Lys Ala Tyr Tyr Lys

420 425 430

Ala Met Thr Lys Ser Ile Asn Lys His Phe Lys Gly Asn Gly Val Ile

435 440 445

Ala Ser Met Glu His Cys Asn Asp Phe Met Phe Leu Gly Thr Glu Ala

450 455 460

Ile Ser Leu Gly Arg Val Gly Asp Asp Phe Trp Cys Thr Asp Pro Ser

465 470 475 480

Gly Asp Pro Asn Gly Thr Phe Trp Leu Gln Gly Cys His Met Val His

485 490 495

Cys Ala Asn Asp Ser Leu Trp Met Gly Asn Phe Ile His Pro Asp Trp

500 505 510

Asp Met Phe Gln Ser Thr His Pro Cys Ala Ala Phe His Ala Ala Ser

515 520 525

Arg Ala Ile Ser Gly Gly Pro Ile Tyr Val Ser Asp Ser Val Gly Lys

530 535 540

His Asn Phe Asp Leu Leu Lys Lys Leu Val Leu Pro Asp Gly Ser Ile

545 550 555 560

Leu Arg Ser Glu Tyr Tyr Ala Leu Pro Thr Arg Asp Cys Leu Phe Glu

565 570 575

Asp Pro Leu His Asn Gly Glu Thr Met Leu Lys Ile Trp Asn Leu Asn

580 585 590

Lys Phe Thr Gly Val Ile Gly Ala Phe Asn Cys Gln Gly Gly Gly Trp

595 600 605

Cys Arg Glu Thr Arg Arg Asn Gln Cys Phe Ser Gln Tyr Ser Lys Arg

610 615 620

Val Thr Ser Lys Thr Asn Pro Lys Asp Ile Glu Trp His Ser Gly Glu

625 630 635 640

Asn Pro Ile Ser Ile Glu Gly Val Lys Thr Phe Ala Leu Tyr Leu Tyr

645 650 655

Gln Ala Lys Lys Leu Ile Leu Ser Lys Pro Ser Gln Asp Leu Asp Ile

/2

660 665 670
Ala Leu Asp Pro Phe Glu Phe Glu Leu Ile Thr Val Ser Pro Val Thr
675 680 685
Lys Leu Ile Gln Thr Ser Leu His Phe Ala Pro Ile Gly Leu Val Asn
690 695 700
Met Leu Asn Thr Ser Gly Ala Ile Gln Ser Val Asp Tyr Asp Asp Asp
705 710 715 720
Leu Ser Ser Val Glu Ile Gly Val Lys Gly Cys Gly Glu Met Arg Val
725 730 735
Phe Ala Ser Lys Lys Pro Arg Ala Cys Arg Ile Asp Gly Glu Asp Val
740 745 750
Gly Phe Lys Tyr Asp Gln Asp Gln Met Val Val Val Gln Val Pro Trp
755 760 765
Pro Ile Asp Ser Ser Ser Gly Gly Ile Ser Val Ile Glu Tyr Leu Phe
770 775 780

Sequence Number: 6
Length of Sequence: 23
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Sequence
TTYTAYCTBA CHGTNCAYCC TCA 23

Sequence Number: 7
Length of Sequence: 23
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Sequence
TTYTAYCTBA CHGTNCAYCC CCA 23

Sequence Number: 8
Length of Sequence: 23
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Sequence
TTYTAYCTBA CHGTNCAYCC ACA 23

Sequence Number: 9
Length of Sequence: 23
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain

Category of Sequence: other nucleic acids, synthetic DNA
Sequence
TTYTAYCTBA CHGTNCAYCC GCA 23

Sequence Number: 10
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 6^{th} and 11^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.
Sequence
GARGGNGTNM GNCAYCTRGT NGAYGG 26

Sequence Number: 11
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 6^{th} and 11^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.
Sequence
GARGGNGTNM GNCAYCTYGT NGAYGG 26

Sequence Number: 12
Length of Sequence: 26
Type of Sequence: nucleic acid /2
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 6^{th} and 11^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.
Sequence
GARGGNGTNM GNCAYTTRGT NGAYGG 26

Sequence Number: 13
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:

Other Information: The 3^{rd} nucleotide N represents inosine, and the other N represent A, G, C, or T.

Sequence

GTNGGNTGYT TYGTNGGYTT YGAYGC 26

Sequence Number: 14
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 3^{rd} nucleotide N represents inosine, and the other N represent A, G, C, or T.

Sequence

GTNGGNTGYT TYGTNGGRTT YGAYGC 26

Sequence Number: 15
Length of Sequence: 29
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 9^{th} and 11^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.

Sequence

TTYGAYGCNT CNGARCCHGA YTCDCGNCA 29

Sequence Number: 16
Length of Sequence: 30
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 9^{th} and 11^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.

Sequence

TTYGAYGCNT CNGARCCHGA YTCDAGYCAY 30

Sequence Number: 17
Length of Sequence: 20
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain

Category of Sequence: other nucleic acids, synthetic DNA
Sequence
GAYCARGAYC TRATGGTNGT 20

Sequence Number: 18
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 6^{th} and 15^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.
Sequence
CCRTCNACYA GRTGNCKNAC NCCYTC 26

Sequence Number: 19
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 6^{th} and 15^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T. /2
Sequence
CCRTCNACRA GRTGNCKNAC NCCYTC 26

Sequence Number: 20
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 6^{th} and 15^{th} nucleotides N represent inosine, and the other N represent A, G, C, or T.
Sequence
CCRTCNACYA TRTGNCKNAC NCCYTC 26

Sequence Number: 21
Length of Sequence: 29
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 3^{rd} and 18^{th} nucleotides N represent inosine, and

the other N represent A, G, C, or T.

Sequence

TGNCGHGART CDGGYTCNGA NGCRTCRAA 29

Sequence Number: 22
Length of Sequence: 30
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Characteristics of Sequence:
Other Information: The 19th nucleotide N represents inosine, and the other N represent A, G, C, or T.

Sequence

RTGRCTHGAR TCDGGYTCNG ANGCRTCRAA 30

Sequence Number: 23
Length of Sequence: 2780
Type of Sequence: nucleic acid
Number of Strands: double strand
Topology: straight chain
Category of Sequence: cDNA to mRNA
Origin
Name of Organism: soybean (Glycine max cv. Clark63)
Characteristics of Sequence
Symbol Expressing Characteristics: CDS
Positions Where Located: 156 … 2405

Sequence

```
TCTTCCATTG GAGGACCATT TCCTCCTGGA ATAGAAATAC TACCACACTT TTCTTTTTTC  60
ACTTCTCTAA GTTGCTAAGT TAATTGCTCC TTCATTTTTT CACTCTTCGT TCTCGCGTAC 120
CCGTGTCACG GTAACTCGTG GTGAAGTGTT CGAAA ATG ACT GTC ACA CCT AAG   173
                                       Met Thr Val Thr Pro Lys
                                        1               5
ATC TCA GTT AAC GAT GGG AAA CTT GTT GTC CAT GGT AAG ACC ATT CTG 221
Ile Ser Val Asn Asp Gly Lys Leu Val Val His Gly Lys Thr Ile Leu
            10                  15                  20
ACT GGA GTG CCA GAC AAC GTT GTG CTG ACT CCA GGT TCT GGA AGG GGT 269
Thr Gly Val Pro Asp Asn Val Val Leu Thr Pro Gly Ser Gly Arg Gly
        25                  30                  35
CTT GTG ACT GGT GCT TTT GTT GGT GCC ACA GCT TCA CAC AGC AAA AGT 317
Leu Val Thr Gly Ala Phe Val Gly Ala Thr Ala Ser His Ser Lys Ser
    40                  45                  50
CTC CAT GTG TTT CCA ATG GGT GTT TTA GAG GGG CTC CGG TTC ATG TGT 365
Leu His Val Phe Pro Met Gly Val Leu Glu Gly Leu Arg Phe Met Cys
55                  60                  65                  70
TGT TTC CGG TTC AAG TTA TGG TGG ATG ACT CAG AGA ATG GGA ACT TGT 413
Cys Phe Arg Phe Lys Leu Trp Trp Met Thr Gln Arg Met Gly Thr Cys
```

```
                      75                  80                  85
GGG AGG GAT GTT CCT CTG GAG ACT CAA TTC ATG CTT ATT GAG AGC AAA     461
Gly Arg Asp Val Pro Leu Glu Thr Gln Phe Met Leu Ile Glu Ser Lys
            90                  95                 100
GAG AGT GAA ACT GAT GGG GAG AAT TCT CCA ATC ATC TAC ACT GTC TTG     509
Glu Ser Glu Thr Asp Gly Glu Asn Ser Pro Ile Ile Tyr Thr Val Leu
        105                 110                 115
CTT CCT CTC CTC GAA GGT CAA TTC CGA GCT GTT CTT CAA GGC AAT GAC     557
Leu Pro Leu Leu Glu Gly Gln Phe Arg Ala Val Leu Gln Gly Asn Asp
    120                 125                 130
AAG AAC GAG ATA GAG ATT TGC CTC GAG AGT GGG GAT AAT GCA GTT GAG     605
Lys Asn Glu Ile Glu Ile Cys Leu Glu Ser Gly Asp Asn Ala Val Glu
135                 140                 145                 150
ACT GAC CAA GGC CTT CAC ATG GTT TAC ATG CAT GCT GGG ACC AAT CCC     653
Thr Asp Gln Gly Leu His Met Val Tyr Met His Ala Gly Thr Asn Pro
                155                 160                 165
TTT GAA GTC ATC AAT CAA GCT GTC AAG GCT GTG GAA AAA CAC ATG CAA     701
Phe Glu Val Ile Asn Gln Ala Val Lys Ala Val Glu Lys His Met Gln
            170                 175                 180
ACT TTT CTT CAT CGT GAG AAG AAA AGG TTG CCA TCT TGT CTT GAC TGG     749
Thr Phe Leu His Arg Glu Lys Lys Arg Leu Pro Ser Cys Leu Asp Trp
        185                 190                 195
TTT GGA TGG TGC ACA TGG GAT GCT TTC TAT ACT GAT GTC ACA GCT GAG     797
Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr Thr Asp Val Thr Ala Glu
    200                 205                 210
GGT GTT GAG GAA GGC CTG AAA AGT CTA TCA CAG GGA GGT ACA CCT CCA     845
Gly Val Glu Glu Gly Leu Lys Ser Leu Ser Gln Gly Gly Thr Pro Pro
215                 220                 225                 230
CGA TTC CTC ATC ATA GAT GAT GGT TGG CAA CAG ATT GAA AAT AAA GCA     893
Arg Phe Leu Ile Ile Asp Asp Gly Trp Gln Gln Ile Glu Asn Lys Ala
                235                 240                 245
AAG GAT GCT ACT GAA TGT TTG GTA CAA GAA GGA GCA CAG TTT GCT ACT     941
Lys Asp Ala Thr Glu Cys Leu Val Gln Glu Gly Ala Gln Phe Ala Thr
            250                 255                 260
AGG TTG ACT GGT ATT AAA GAG AAT ACT AAA TTT CAA AAG AAA TTA CAG     989
Arg Leu Thr Gly Ile Lys Glu Asn Thr Lys Phe Gln Lys Lys Leu Gln
        265                 270                 275
AAC AAT GAG CAG ATG TCA GGT CTG AAG CAT CTA GTA CAT GGA GCA AAG    1037
Asn Asn Glu Gln Met Ser Gly Leu Lys His Leu Val His Gly Ala Lys
    280                 285                 290
CAG CAT CAC AAT GTG AAA AAT GTA TAT GTA TGG CAT GCA CTA GCT GGT    1085
Gln His His Asn Val Lys Asn Val Tyr Val Trp His Ala Leu Ala Gly
295                 300                 305                 310
TAT TGG GGT GGA GTG AAG CCA GCA GCA ACC GGC ATG GAA CAT TAT GAC    1133
Tyr Trp Gly Gly Val Lys Pro Ala Ala Thr Gly Met Glu His Tyr Asp
                315                 320                 325
ACT GCC TTG GCA TAT CCA GTG CAG TCA CCA GGC GTG CTA GGA AAC CAA    1181
Thr Ala Leu Ala Tyr Pro Val Gln Ser Pro Gly Val Leu Gly Asn Gln
            330                 335                 340
CCA GAC ATT GTC ATG GAC AGC TTG GCT GTA CAT GGC CTT GGC CTA GTG    1229
```

/3

Pro Asp Ile Val Met Asp Ser Leu Ala Val His Gly Leu Gly Leu Val
345 350 355

CAC CCA AAG AAG GTT TTC AAT TTC TAC AAC GAG CTC CAT GCT TAC TTA 1277
His Pro Lys Lys Val Phe Asn Phe Tyr Asn Glu Leu His Ala Tyr Leu
360 365 370

GCT TCT TGT GGA GTA GAT GGA GTG AAG GTT GAT GTG CAG AAC ATT ATT 1325
Ala Ser Cys Gly Val Asp Gly Val Lys Val Asp Val Gln Asn Ile Ile
375 380 385 390

GAG ACC CTT GGT GCG GGA CAT GGT GGC CGA GTG TCA CTT ACT CGC AGC 1373
Glu Thr Leu Gly Ala Gly His Gly Gly Arg Val Ser Leu Thr Arg Ser
395 400 405

TAT CAT CAC GCG CTT GAG GCT TCC ATT GCT AGC AAT TTT ACT GAT AAC 1421
Tyr His His Ala Leu Glu Ala Ser Ile Ala Ser Asn Phe Thr Asp Asn
410 415 420

GGA TGC ATT GCG TGT ATG TGT CAC AAC ACT GAT GGA CTT TAT AGT GCT 1469
Gly Cys Ile Ala Cys Met Cys His Asn Thr Asp Gly Leu Tyr Ser Ala
425 430 435

AAG CAG ACT GCT ATT GTG AGA GCT TCT GAT GAT TTT TAC CCT CGT GAT 1517
Lys Gln Thr Ala Ile Val Arg Ala Ser Asp Asp Phe Tyr Pro Arg Asp
440 445 450

CCT GCT TCC CAT ACC ATC CAT ATT TCT TCT GTT GCA TAC AAC TCA CTA 1565
Pro Ala Ser His Thr Ile His Ile Ser Ser Val Ala Tyr Asn Ser Leu
455 460 465 470

TTC CTT GGA GAA TTC ATG CAA CCT GAC TGG GAC ATG TTT CAT AGT TTA 1613
Phe Leu Gly Glu Phe Met Gln Pro Asp Trp Asp Met Phe His Ser Leu
475 480 485

CAC CCA GCA GCA GAT TAT CAT GCT GCA GCT CGT GCA ATT GGT GGA TGT 1661
His Pro Ala Ala Asp Tyr His Ala Ala Ala Arg Ala Ile Gly Gly Cys
490 495 500

CCT ATT TAT GTT AGT GAC AAG CCA GGC AAT CAC AAT TTT GAT CTT CTT 1709
Pro Ile Tyr Val Ser Asp Lys Pro Gly Asn His Asn Phe Asp Leu Leu
505 510 515

AAG AAG CTG GTT CTC CCG GAT GGT TCG GTT CTC CGT GCT CAG TTA CCT 1757
Lys Lys Leu Val Leu Pro Asp Gly Ser Val Leu Arg Ala Gln Leu Pro
520 525 530

GGC AGG CCA ACT CGT GAT TCT CTA TTT GTG GAT CCA GCC AGA GAT AGG 1805
Gly Arg Pro Thr Arg Asp Ser Leu Phe Val Asp Pro Ala Arg Asp Arg
535 540 545 550

ACT AGC TTG CTC AAA ATA TGG AAC CTG AAC AAA TGC TCT GGA GTT GTT 1853
Thr Ser Leu Leu Lys Ile Trp Asn Leu Asn Lys Cys Ser Gly Val Val
555 560 565

GGT GTA TTT AAC TGC CAA GGT GCT GGA TGG TGC AAG ATA GAG AAG AAA 1901
Gly Val Phe Asn Cys Gln Gly Ala Gly Trp Cys Lys Ile Glu Lys Lys
570 575 580

ACC CGC ATC CAT GAT ACA TCT CCT GGT ACA CTC ACC GCC TCT GTC TGC 1949
Thr Arg Ile His Asp Thr Ser Pro Gly Thr Leu Thr Ala Ser Val Cys
585 590 595

GCC TCT GAT GTT GAC CTC ATC ACA CAA GTA GCA GGT GCT GAA TGG CTT 1997
Ala Ser Asp Val Asp Leu Ile Thr Gln Val Ala Gly Ala Glu Trp Leu
600 605 610

GGA GAT ACA ATT GTT TAT GCT TAC AGA TCA GGT GAG GTG ATT CGG CTA 2045
Gly Asp Thr Ile Val Tyr Ala Tyr Arg Ser Gly Glu Val Ile Arg Leu
615 620 625 630
CCA AAA GGG GTT TCA ATT CCA GTG ACA CTA AAA GTT CTG GAG TTT GAG 2093
Pro Lys Gly Val Ser Ile Pro Val Thr Leu Lys Val Leu Glu Phe Glu
635 640 645
CTT TTC CAC TTC TGT CCA ATC CAA GAA ATA GCT CCA AGT ATA TCA TTT 2141
Leu Phe His Phe Cys Pro Ile Gln Glu Ile Ala Pro Ser Ile Ser Phe
650 655 660
GCA GCA ATA GGG CTA CTG GAT ATG TTC AAC ACT GGA GGA GCA GTG GAG 2189
Ala Ala Ile Gly Leu Leu Asp Met Phe Asn Thr Gly Gly Ala Val Glu
665 670 675
CAG GTT GAG ATT CAT AAC CGA GCA GCA ACG AAA ACA ATA GCT CTT AGT 2237
Gln Val Glu Ile His Asn Arg Ala Ala Thr Lys Thr Ile Ala Leu Ser
680 685 690
GTA AGG GGA AGA GGC AGA TTT GGA GTT TAC TCC TCC CAG AGA CCA CTG 2285
Val Arg Gly Arg Gly Arg Phe Gly Val Tyr Ser Ser Gln Arg Pro Leu
695 700 705 710
AAG TGT GTG GTA GGT GGC GCT GAA ACC GAC TTC AAC TAT GAC TCA GAG 2333
Lys Cys Val Val Gly Gly Ala Glu Thr Asp Phe Asn Tyr Asp Ser Glu
715 720 725
ACC GGG TTG ACA ACC TTC TCC ATT CCA GTT TCT CCA GAG GAG ATG TAC 2381
Thr Gly Leu Thr Thr Phe Ser Ile Pro Val Ser Pro Glu Glu Met Tyr
730 735 740
AGA TGG TCA ATA GAG ATC CAA GTT TGAGTCCTTT TTAAGACTTG GTGTTTGATG 2435
Arg Trp Ser Ile Glu Ile Gln Val
745 750
CATTGTTGTA TCAGGAGAAG GGTTTTGTTG TAATTAAGCA TTGAGGGAAT TGTTGGAGTC 2495
AGGCAGAGAG AGAGGGGGGA GGTTTGTTGT AAGACACCTA GTATTAGTAT CATGTAGTGG 2555
AGAAAAAGGG TTGTTGATCC TAATAGCTAG ACAAGGCATG TTGTAGTAGT CATGGGGTGG 2615
GGAAGTCCTT TTGTTGTAGC ATGTAATTTG GTTTAGACTT GTAGTATGTC ATCAATTAGA 2675
TGGATAAAGA GAGAATATTG TTATCTACCC GAGGATGTAA CAATGTTTGT TTCTCTGAAT 2735
AAAAAGTTCA CATCTGTCTT TTGGAATAAT AAAAAAAAAA AAAAA 2780

Sequence Number: 24
Length of Sequence: 750
Type of Sequence: amino acid
Topology: straight chain
Category of Sequence: protein

Sequence

Met Thr Val Thr Pro Lys Ile Ser Val Asn Asp Gly Lys Leu Val Val
1 5 10 15
His Gly Lys Thr Ile Leu Thr Gly Val Pro Asp Asn Val Val Leu Thr
20 25 30
Pro Gly Ser Gly Arg Gly Leu Val Thr Gly Ala Phe Val Gly Ala Thr
35 40 45
Ala Ser His Ser Lys Ser Leu His Val Phe Pro Met Gly Val Leu Glu
50 55 60
Gly Leu Arg Phe Met Cys Cys Phe Arg Phe Lys Leu Trp Trp Met Thr
65 70 75 80
Gln Arg Met Gly Thr Cys Gly Arg Asp Val Pro Leu Glu Thr Gln Phe
85 90 95
Met Leu Ile Glu Ser Lys Glu Ser Glu Thr Asp Gly Glu Asn Ser Pro

```
                100                 105                 110
Ile Ile Tyr Thr Val Leu Leu Pro Leu Leu Glu Gly Gln Phe Arg Ala
        115                 120                 125
Val Leu Gln Gly Asn Asp Lys Asn Glu Ile Glu Ile Cys Leu Glu Ser
    130                 135                 140
Gly Asp Asn Ala Val Glu Thr Asp Gln Gly Leu His Met Val Tyr Met
145                 150                 155                 160
His Ala Gly Thr Asn Pro Phe Glu Val Ile Asn Gln Ala Val Lys Ala
                165                 170                 175
Val Glu Lys His Met Gln Thr Phe Leu His Arg Glu Lys Lys Arg Leu
            180                 185                 190
Pro Ser Cys Leu Asp Trp Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr
        195                 200                 205
Thr Asp Val Thr Ala Glu Gly Val Glu Glu Gly Leu Lys Ser Leu Ser
    210                 215                 220
Gln Gly Gly Thr Pro Pro Arg Phe Leu Ile Ile Asp Asp Gly Trp Gln
225                 230                 235                 240
Gln Ile Glu Asn Lys Ala Lys Asp Ala Thr Glu Cys Leu Val Gln Glu
                245                 250                 255
Gly Ala Gln Phe Ala Thr Arg Leu Thr Gly Ile Lys Glu Asn Thr Lys
            260                 265                 270
Phe Gln Lys Lys Leu Gln Asn Asn Glu Gln Met Ser Gly Leu Lys His
        275                 280                 285
Leu Val His Gly Ala Lys Gln His His Asn Val Lys Asn Val Tyr Val
    290                 295                 300
Trp His Ala Leu Ala Gly Tyr Trp Gly Gly Val Lys Pro Ala Ala Thr
305                 310                 315                 320
Gly Met Glu His Tyr Asp Thr Ala Leu Ala Tyr Pro Val Gln Ser Pro
                325                 330                 335
Gly Val Leu Gly Asn Gln Pro Asp Ile Val Met Asp Ser Leu Ala Val
            340                 345                 350
His Gly Leu Gly Leu Val His Pro Lys Lys Val Phe Asn Phe Tyr Asn
        355                 360                 365
Glu Leu His Ala Tyr Leu Ala Ser Cys Gly Val Asp Gly Val Lys Val
    370                 375                 380
Asp Val Gln Asn Ile Ile Glu Thr Leu Gly Ala Gly His Gly Gly Arg
385                 390                 395                 400
Val Ser Leu Thr Arg Ser Tyr His His Ala Leu Glu Ala Ser Ile Ala
                405                 410                 415
Ser Asn Phe Thr Asp Asn Gly Cys Ile Ala Cys Met Cys His Asn Thr
            420                 425                 430
Asp Gly Leu Tyr Ser Ala Lys Gln Thr Ala Ile Val Arg Ala Ser Asp
        435                 440                 445
Asp Phe Tyr Pro Arg Asp Pro Ala Ser His Thr Ile His Ile Ser Ser
    450                 455                 460
Val Ala Tyr Asn Ser Leu Phe Leu Gly Glu Phe Met Gln Pro Asp Trp
465                 470                 475                 480
Asp Met Phe His Ser Leu His Pro Ala Ala Asp Tyr His Ala Ala Ala
                485                 490                 495
Arg Ala Ile Gly Gly Cys Pro Ile Tyr Val Ser Asp Lys Pro Gly Asn
```

/3

500 505 510
His Asn Phe Asp Leu Leu Lys Lys Leu Val Leu Pro Asp Gly Ser Val
515 520 525
Leu Arg Ala Gln Leu Pro Gly Arg Pro Thr Arg Asp Ser Leu Phe Val
530 535 540
Asp Pro Ala Arg Asp Arg Thr Ser Leu Leu Lys Ile Trp Asn Leu Asn
545 550 555 560
Lys Cys Ser Gly Val Val Gly Val Phe Asn Cys Gln Gly Ala Gly Trp
565 570 575
Cys Lys Ile Glu Lys Lys Thr Arg Ile His Asp Thr Ser Pro Gly Thr
580 585 590
Leu Thr Ala Ser Val Cys Ala Ser Asp Val Asp Leu Ile Thr Gln Val
595 600 605
Ala Gly Ala Glu Trp Leu Gly Asp Thr Ile Val Tyr Ala Tyr Arg Ser
610 615 620
Gly Glu Val Ile Arg Leu Pro Lys Gly Val Ser Ile Pro Val Thr Leu
625 630 635 640
Lys Val Leu Glu Phe Glu Leu Phe His Phe Cys Pro Ile Gln Glu Ile
645 650 655
Ala Pro Ser Ile Ser Phe Ala Ala Ile Gly Leu Leu Asp Met Phe Asn
660 665 670
Thr Gly Gly Ala Val Glu Gln Val Glu Ile His Asn Arg Ala Ala Thr
675 680 685
Lys Thr Ile Ala Leu Ser Val Arg Gly Arg Gly Arg Phe Gly Val Tyr
690 695 700
Ser Ser Gln Arg Pro Leu Lys Cys Val Val Gly Gly Ala Glu Thr Asp
705 710 715 720
Phe Asn Tyr Asp Ser Glu Thr Gly Leu Thr Thr Phe Ser Ile Pro Val
725 730 735
Ser Pro Glu Glu Met Tyr Arg Trp Ser Ile Glu Ile Gln Val
740 745 750

Sequence Number: 25
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Sequence
ATSCAVCCTG ACTGGGATAT GTTCCA 26

Sequence Number: 26
Length of Sequence: 26
Type of Sequence: nucleic acid
Number of Strands: single strand
Topology: straight chain
Category of Sequence: other nucleic acids, synthetic DNA
Sequence
CGAAGGAYYG AWCCATCAGG AARHAM 26

Sequence Number: 27
Length of Sequence: 253
Type of Sequence: nucleic acid
Number of Strands: double strand
Topology: straight chain
Category of Sequence: cDNA to mRNA
Origin
Name of Organism: Arabidopsis (Arabidopsis thaliana)

Sequence

```
GGCTTATGCA ACCTGACTGG GAATGTTCCA TAGTCTACAC CCAACTGCAG AGTACCATGC  60
TGCAGCGCGT GCAGTGGGTG GATGCGCAAT CTATGTCAGT GATAAGCCAG GCAACCACAA 120
CTTTGATCTA TTGAGGAAGC TGGTTCTTCC TGATGGTTCA GTTCTTCGGG CTAAGCTCCC 180
GGGTAGGCCT ACCCGTGACT GCTTATTCGC TGATCCAGCT AGAGATGGGA TCAGCTTGCT 240

CAAGATCTGG AAC                                                    253
```

/3

Sequence Number: 28
Length of Sequence: 10
Type of Sequence: amino acid
Topology: straight chain
Category of Sequence: peptide

Sequence

```
Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr
 1               5                  10
```

Sequence Number: 29
Length of Sequence: 13
Type of Sequence: amino acid
Topology: straight chain
Category of Sequence: peptide
Other Information: The 8^{th} amino acid Xaa represents Ala or Cys.

Sequence

```
Val Tyr Val Trp His Ala Leu Xaa Gly Tyr Trp Gly Gly
 1               5                  10
```

Sequence Number: 28
Length of Sequence: 10
Type of Sequence: amino acid
Topology: straight chain
Category of Sequence: peptide

Sequence

```
His Asn Phe Asp Leu Leu Lys Lys Leu Val Leu Pro Asp Gly Ser
 1               5                  10                  15
```

[Brief Explanation of the Drawings]

[Fig. 1] Graph showing the correlation between production of raffinose by raffinose synthesis and reaction time.

[Fig. 2] Schematic diagram showing results of SDS-polyacrylamide gel electrophoresis of raffinose synthetase. M represents a molecular weight marker, and S represents a sample containing raffinose synthetase. Numbers represent molecular weight (kDa).

[Fig. 3] Graph showing the effect of reaction temperature on raffinose synthetase activity.

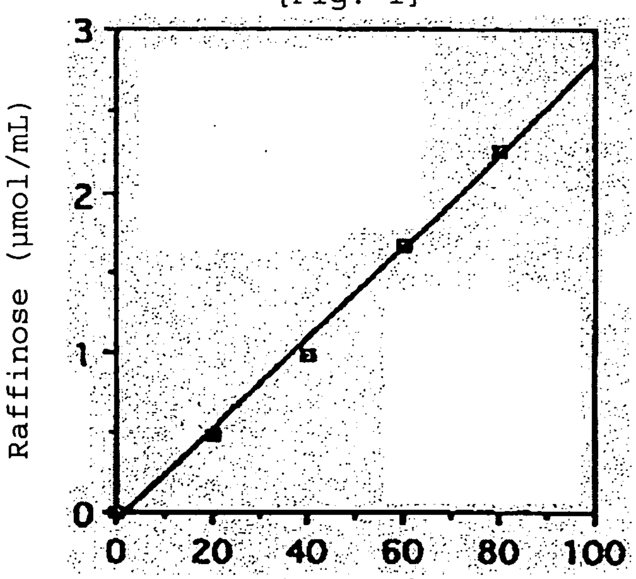
[Fig. 4] Graph showing the effect of pH on raffinose synthetase activity.

[Fig. 5] Graph showing the effect of myoinositol on raffinose synthetase activity.

[Fig. 6] Graph showing the stable pH range of raffinose synthetase.

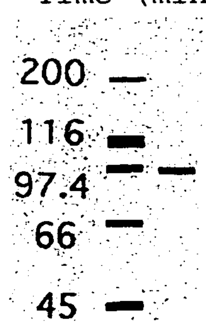
[Fig. 7] Diagram showing the correlation between synthesis primers and amino acid sequences of peptides. R represents A or G, Y represents C or T, M represents A or C, K represents G or T, D represents G, A, or T, H represents A, T, or C, B represents T or C, and N represents G, A, G, C, or inosine.

[Fig. 1]

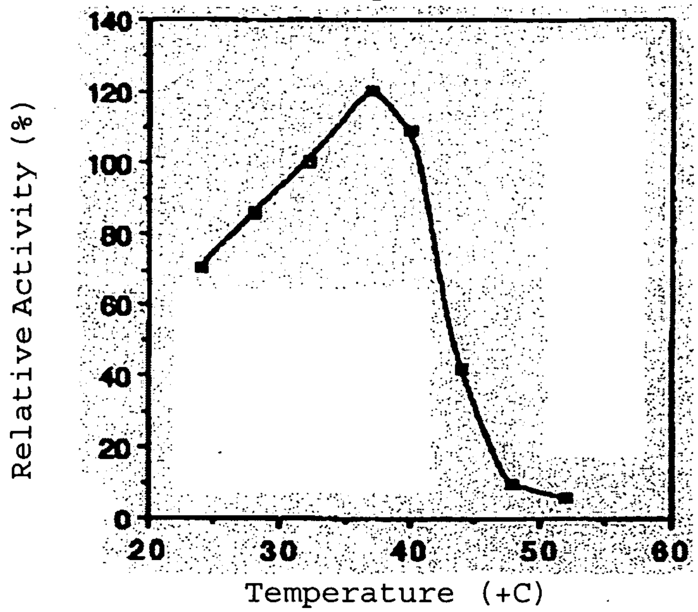


[Fig. 2]

Time (min)

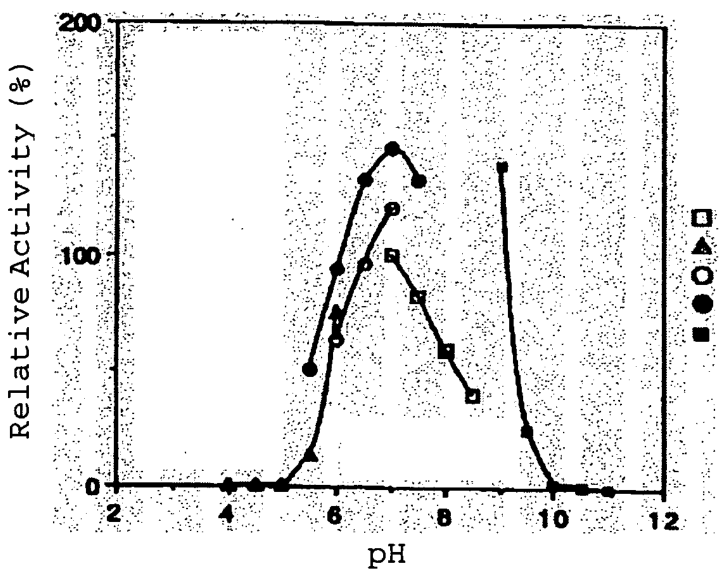


[Fig. 3]



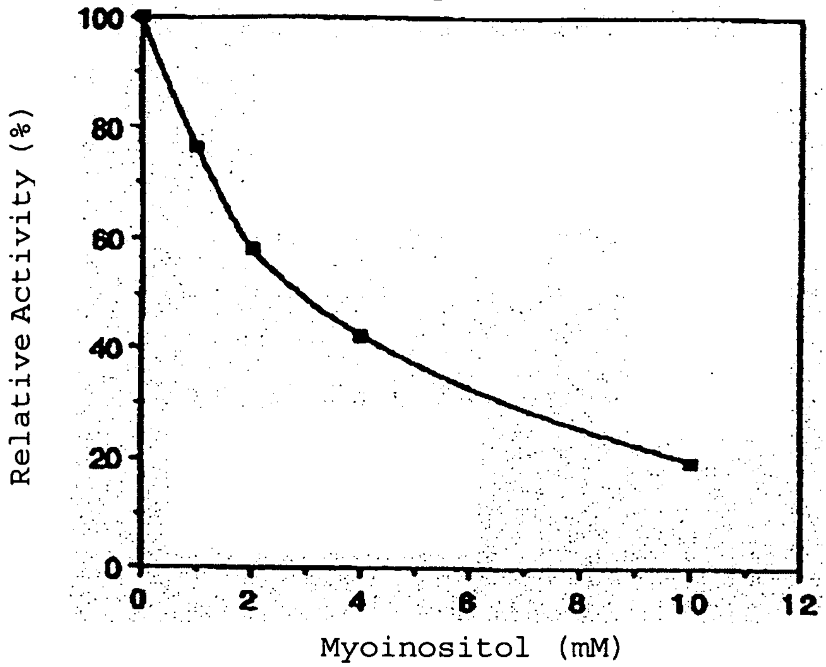
140
120
100
80
60
40
20
0
20
30
40
50
60
Relative Activity (%)
Temperature (+C)

[Fig. 4]

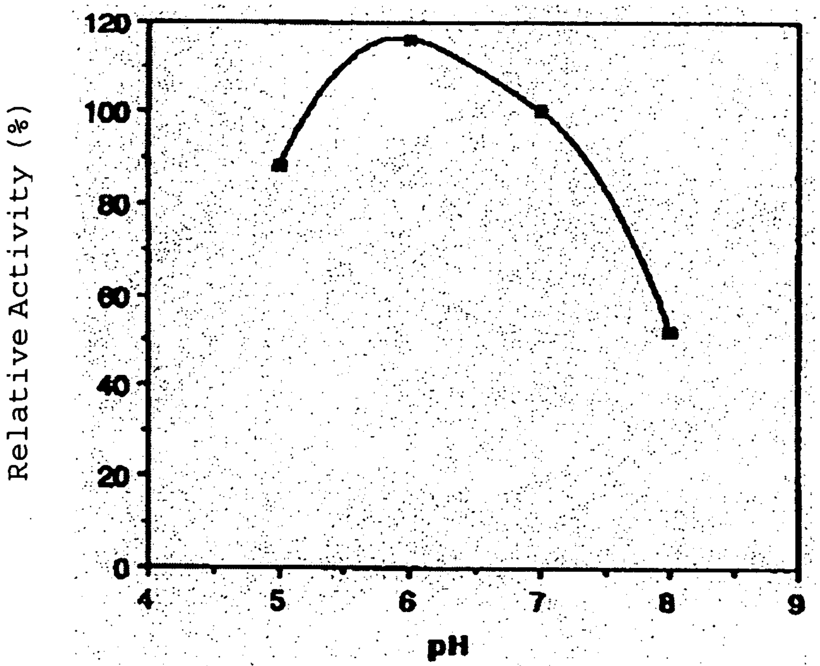


200
100
0
2
4
6
8
10
12
Relative Activity (%)
pH

[Fig. 5]


100
80
60
40
20
0
0
2
4
6
8
10
12
Relative Activity (%)
Myoinositol (mM)

[Fig. 6]


120
100
80
60
40
20
0
4
5
6
7
8
9
Relative Activity (%)
pH

[Fig. 7]

Sequence Number 6 A1 5' - TTY TAY CTB ACH GTN CAY CCT CA -3'

Sequence Number 7 A2 5' - TTY TAY CTB ACH GTN CAY CCC CA -3' B1 5' - GAR GGN GTN MGN CAY CTR GTN GAY GG - Sequence Number 10

Sequence Number 8 A3 5' - TTY TAY CTB ACH GTN CAY CCA CA -3' B2 5' - GAR GGN GTN MGN CAY CTY GTN GAY GG - Sequence Number 11

Sequence Number 9 A4 5' - TTY TAY CTB ACH GTN CAY CCG CA -3' B3 5' - GAR GGN GTN MGN CAY TTR GTN GAY GG - Sequence Number 12

Phe Gly Trp Cys Thr Trp Asp Ala Phe Tyr Leu Thr Val His Pro Gln Gly Val Ile Glu Gly Val Arg His Leu Val Asp Gly Gly Cys

Sequence Number 1

Sequence Number 18 3' - CTY CCN CAN KCI GTR GAY CAI CTR CC -5' B'1

Sequence Number 19 3' - CTY CCN CAN KCI GTR GAR CAI CTR CC -5' B'2

Sequence Number 20 3' - CTY CCN CAN KCI GTR TAY CAI CTR CC -5' B'3

D1 5' - TTY GAY GCN TCN GAR CCH GAY TCD CGN CA -3' Sequence Number 15

D2 5' - TTY GAY GCN TCN GAR CCH GAY TCD AGY CAY -3'

C1 5' - GTN GGN TGY TTY GTN GGY TTY GAY GC -3' Sequence Number 13

C2 5' - GTN GGN TGY TTY GTN GGR TTY GAY GC -3' Sequence Number 14

Pro Val Ser Val Gly Cys Phe Val Gly Phe Asp Ala Ser Glu Pro Asp Ser Arg His

Sequence Number 2

3' - AAR CTR CGN AGI CTY GGD CTR AGH GCI GT -5' D'1 Sequence Number 21

3' - AAR CTR CGN AGI CTY GGD CTR AGH TCR GTR -5' D'2 Sequence Number 22

E 5' - GAY CAR GAY CTR ATG GTN GT -3' Sequence Number 17

Tyr Asp Gln Asp Gln Met Val Val Val Gln Val Pro Trp Pro

Sequence Number 3

Continuation of 3. NOTE: Continuation of 3. NOTE: (a) addition of the limitations "growing plants under drought conditions " and "growin plants under high salt conditions" raises new issues that would require further consideration under 35 USC 112, 1st paragraph with respect to growing plants under drought or high salt conditions as a means of increasing the drought resistance and resistance to high salt conditions in plants; addition of new claims directed to increasing the raffinose synthase activity of plants raises new issues that woul require further consideration under 35 USC 112, 1st paragraph and 35 USC 102/103.

Continuation of 5. Applicant's reply would overcome the following rejection(s): addition of the proposed limitations "growing plants under drought conditions" and "growing plants under high salt conditions" would, if entered, overcome the rejection of claims 28-47 under 35 U.S.C. 102(a) as being anticipated by EP 0 994 186 A1 (AJINOMOTO CO. INC., 19.04.2000), the rejection of claims 28-31 and 38-41 under 35 U.S.C. 102(b) as being anticipated by EP 0 849 359 A2 (SUMITOMO CHEMICAL CO, 24.06.1998), and the rejection of Claims 28-47 under 35 U.S.C. 102(b) as being anticipated by JP411123080-A (AJINOMOTO CO. INC., May 11, 1999).

Continuation of 11. does NOT place the application in condition for allowance because: Applicants' request for reconsideration relies on entry of the proposed amendments, which proposed amendments will not be entered as set forth above.